

STIC-Biotech/ChemLib

90905

From:
Sent:
To:
Subject:

Chan, Christina
Monday, April 07, 2003 4:17 PM
Nguyen, Quang (AU1632); STIC-Biotech/ChemLib
RE: RUSH sequence search request for 09/761640

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

CRFE

-----Original Message-----

From: **Nguyen, Quang (AU1632)**
Sent: Monday, April 07, 2003 4:14 PM
To: Chan, Christina
Subject: RUSH sequence search request for 09/761640

I would like to request for a RUSH sequence search for the above application, because I need a count for this bi-week.

Please search:

A nucleic acid encoding the amino acid of SEQ ID NO:4;
A nucleic acid of SEQ ID NO:1; and
A nucleic acid of SEQ ID NO:3;
against commercial, issued and pending US patent application databases.

I am in AU1636, my mailbox is in CM1-11E12.

THANK YOU.

Edward Hart
Technical Info. Specialist
STIC/Biotech
April 2002 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/8/03
Date Completed: 4/14/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: 9
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

revised
NA

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 1#
WWW/Internet: _____
Other (specify): _____


```

Db 1681 CTTCTCTCCACGAGCTCTTCATGAGAGAGCTCTGACGCCCTTCCACAGCTTGCAAG 1740
QY 1741 GACCAAGGAGAGCCACAGAGGTGAGAGGGGGCTTACGCTGCTGGAAGTCCCGCACTC 1800
Db 1741 GACCAAGGAGAGCCACAGAGGTGAGAGGGGGCTTACGCTGCTGGAAGTCCCGCACTC 1800
QY 1801 AGTGTTACCTCCACAGAGGACATGCGTGTGGCCAAACCGGAGCCAGGCTTCCAGAGCA 1860
Db 1801 AGTGTTACCTCCACAGAGGACATGCGTGTGGCCAAACCGGAGCCAGGCTTCCAGAGCA 1860
QY 1861 GGAGCAGGGGAGGGGAGGGGAGAGAGAGCCCTGCAATTTCTCTACAGCCAGGTTCCG 1920
Db 1861 GGAGCAGGGGAGGGGAGGGGAGAGAGAGCCCTGCAATTTCTCTACAGCCAGGTTCCG 1920
QY 1921 GAAGGTGTGAGAGAGGACGACCTGTCATGAGTGTGAGAGAGAGGGGAGGCTGAGCCCT 1980
Db 1921 GAAGGTGTGAGAGAGGACGACCTGTCATGAGTGTGAGAGAGAGGGGAGGCTGAGCCCT 1980
QY 1981 CACACATGCCACGCTCCCTGACACTGAAAGAGATCCACAACTCTTGGAGAAACACCC 2040
Db 1981 CACACATGCCACGCTCCCTGACACTGAAAGAGATCCACAACTCTTGGAGAAACACCC 2040
QY 2041 TCACGTGTGTGCGGACACATTCCTCTCAGCTCCGCCCATACCCGTCTACAGCCTC 2100
Db 2041 TCACGTGTGTGCGGACACATTCCTCTCAGCTCCGCCCATACCCGTCTACAGCCTC 2100
QY 2101 ACCGTCCACCCCTGTCACCTAGAGGCTCAGCTCCACCCCTGTCACCTACAGCCTCCTCC 2160
Db 2101 ACCGTCCACCCCTGTCACCTAGAGGCTCAGCTCCACCCCTGTCACCTACAGCCTCCTCC 2160
QY 2161 TACAGCCTTAAGTCCACAGGCGCATGTCTGCTTCACAGGGGCTCAGAGCTTCTAATGG 2220
Db 2161 TACAGCCTTAAGTCCACAGGCGCATGTCTGCTTCACAGGGGCTCAGAGCTTCTAATGG 2220
QY 2221 GATGTGTAGAGGAGCTGAGATCCTTTGGGGGCAACAGACACCTAGTTTCATTCTCAA 2280
Db 2221 GATGTGTAGAGGAGCTGAGATCCTTTGGGGGCAACAGACACCTAGTTTCATTCTCAA 2280
QY 2281 CTCTAGCCCTGACACACTGCTGTGGGACGGAATGAAACAGAGCTTCCCGGAAAAAG 2340
Db 2281 CTCTAGCCCTGACACACTGCTGTGGGACGGAATGAAACAGAGCTTCCCGGAAAAAG 2340
QY 2341 GGTACAGGCTCCACACCCCGCCCTCCCTGACCTCTGTCTCTCTCCAGTTCAATTCCT 2400
Db 2341 GGTACAGGCTCCACACCCCGCCCTCCCTGACCTCTGTCTCTCTCCAGTTCAATTCCT 2400
QY 2401 GGAACCAAGCCAGGCGCAAGCAAGTGGCCGCCAAAGGAGGAGATCTCAGGCCCCA 2460
Db 2401 GGAACCAAGCCAGGCGCAAGCAAGTGGCCGCCAAAGGAGGAGATCTCAGGCCCCA 2460
QY 2461 GCGCGGGAGGCTGGAAGGGGTGGCAGATGCTCTCCATCCACCTCCACCGGTCAGG 2520
Db 2461 GCGCGGGAGGCTGGAAGGGGTGGCAGATGCTCTCCATCCACCTCCACCGGTCAGG 2520
QY 2521 TCTTGTGCTGTGCTCCAGACCTCTGTGACACACGCGCAATACAGAGGAGCCAGGCA 2580
Db 2521 TCTTGTGCTGTGCTCCAGACCTCTGTGACACACGCGCAATACAGAGGAGCCAGGCA 2580
QY 2581 GAGATAGTCTTCTTTTGTCTCTTCTGTGCGCTCTGAGTATGATTTTTCATAGCCTTACA 2640
Db 2581 GAGATAGTCTTCTTTTGTCTCTTCTGTGCGCTCTGAGTATGATTTTTCATAGCCTTACA 2640
QY 2641 GTATCTGGCTTTTGTACTGAGAAATTAATTAATTTTCTAATAAAAAAAAAAAAAAAAAA 2700
Db 2641 GTATCTGGCTTTTGTACTGAGAAATTAATTAATTTTCTAATAAAAAAAAAAAAAAAAAA 2700
QY 2701 AAAA 2704
Db 2701 AAAA 2704

```

```

AB073252
ID AB073252 standard; cdna; 2704 bp.
AC AC
AC AB073252;
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 1 cDNA sequence spq ID NO:7.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;
KW SNP; gene; ss.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT 5'UTR 1..93
FT CDS /*tag= a
FT 94..1509
FT /*tag= b
FT /product= "MAP kinase phosphatase splice form 1"
FT variation replace(577,A)
FT /*tag= c
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /*tag= d
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT 3'UTR 1510..2704
FT variation /*tag= e
FT /*tag= (2641,A)
FT /*tag= f
FT /*standard_name= "single nucleotide polymorphism (SNP)"
XX
XX W0200242436-A2.
XX
XX 30-MAY-2002.
XX
XX 07-NOV-2001; 2001WO-US42995.
XX
XX 20-NOV-2000; 2000US-0715177.
XX 18-JAN-2001; 2001US-0761640.
XX
XX (PEKE ) PE CORP NY.
XX
XX WPI: 2002-575237/61.
XX P-PSDB; ABP51653.
XX
XX Novel isolated human phosphatase peptide useful for treating disorder
XX characterized by absence of, inappropriate or unwanted expression of
XX the phosphatase protein, and as immunogens to raise antibodies
XX
XX Claim 1; Fig 3A; 85bp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
XX (I). (I) can be used for identifying a modulator of (I) by contacting
XX (I) with an agent and determining if the agent has modulated the
XX function or activity of (I). (I) is useful for identifying an agent that
XX binds to (I), by contacting (I) with an agent and assaying the contacted
XX mixture to determine whether a complex is formed with the agent bound
XX (I). The human phosphatases from the present invention are mitogen
XX activated protein (MAP) kinase phosphatases. These human MAP kinase
XX phosphatases are located on chromosome 11. (I) and the polynucleotide
XX sequences encoding (I) can be used in gene therapy. The present sequence
XX encodes human MAP kinase phosphatase splice form 1 from the present
XX invention.
XX
XX Sequence 2704 bp; 569 A; 874 C; 794 G; 467 T; 0 other;
XX
Query Match 100.0%; Score 2704; DB 24; Length 2704;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CGTCTTCTGTCCTGCGGGTCCAGGACTGTCGCGGGTTGAGGGAAGGGCCGTGCC 60
 Db 1 CGTCTTCTGTCCTGCGGGTCCAGGACTGTCGCGGGTTGAGGGAAGGGCCGTGCC 60
 QY 61 CGGTGCCAGCCAGGTGTCGCGGCTGGCTCATGGCCCTGTGTCACAGTGCCTTGC 120
 Db 61 CGGTGCCAGCCAGGTGTCGCGGCTGGCTCATGGCCCTGTGTCACAGTGCCTTGC 120
 QY 121 CCGCCGGGCGCGGGCTCCAGCGCCGTGGGGCCCGGGAGCCAGGGGCTCCAGGAG 180
 Db 121 CCGCCGGGCGCGGGCTCCAGCGCCGTGGGGCCCGGGAGCCAGGGGCTCCAGGAG 180
 QY 181 AGTCGACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 Db 181 AGTCGACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 QY 241 GATGAGAGGAGCAATATGATGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 Db 241 GATGAGAGGAGCAATATGATGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 QY 301 AGTGAAGAGAGGTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 Db 301 AGTGAAGAGAGGTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 QY 361 AAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 420
 Db 361 AAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 420
 QY 421 GATATCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 Db 421 GATATCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 QY 481 GTAGTTTCTACAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 Db 481 GTAGTTTCTACAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 QY 541 GATTTCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 Db 541 GATTTCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
 QY 601 GATTTCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 Db 601 GATTTCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 QY 661 CGGATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 Db 661 CGGATCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 QY 721 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 Db 721 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 QY 781 AGCCACTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
 Db 781 AGCCACTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
 QY 841 GCGGACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 Db 841 GCGGACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 QY 901 CAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 Db 901 CAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 QY 961 AGTGTACTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 Db 961 AGTGTACTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
 QY 1021 CAGTACCGTACTTCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
 Db 1021 CAGTACCGTACTTCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080

QY 1081 TCCGCGATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGAGGAGGAGGAGGAGGAG 1140
 Db 1081 TCCGCGATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGAGGAGGAGGAGGAGGAG 1140
 QY 1141 CTGCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
 Db 1141 CTGCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
 QY 1201 CCGAGCGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1260
 Db 1201 CCGAGCGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1260
 QY 1261 CCGAGCGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1320
 Db 1261 CCGAGCGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 1320
 QY 1321 CTGTCGCTGCAAGAGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
 Db 1321 CTGTCGCTGCAAGAGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
 QY 1381 AAGCAGTACGAAATGACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 Db 1381 AAGCAGTACGAAATGACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
 QY 1441 GCGCGCCCAACCTGCTGCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
 Db 1441 GCGCGCCCAACCTGCTGCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
 QY 1501 AGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
 Db 1501 AGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
 QY 1561 GAAAGAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
 Db 1561 GAAAGAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
 QY 1621 TCTTCTGAGAGGCTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
 Db 1621 TCTTCTGAGAGGCTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
 QY 1681 CTTCCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
 Db 1681 CTTCCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
 QY 1741 GACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
 Db 1741 GACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
 QY 1801 AGTGTACTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
 Db 1801 AGTGTACTTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
 QY 1861 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
 Db 1861 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
 QY 1921 GAAGGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
 Db 1921 GAAGGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
 QY 1981 CACACATGCGCCAGGCTGCTGACACTGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 2040
 Db 1981 CACACATGCGCCAGGCTGCTGACACTGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 2040
 QY 2041 TCAGGCTGTGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
 Db 2041 TCAGGCTGTGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
 QY 2101 ACCTCCAGCCCTGTGACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
 Db 2101 ACCTCCAGCCCTGTGACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
 QY 2161 TACAGCCTTAAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220

|||||
Db 2161 TACAGCCCTTAAGTCCAGAGCCCATGCTGCTGCTCAAGGCGTCAAGACTTCTAACTGG 2220
Oy 2221 GATGGGAGGAGGAGCAAGTACCTTTGGGGGCAAGAGACCTTATTCATTCTCAA 2280
Db 2221 GATGGGAGGAGGAGCAAGTACCTTTGGGGGCAAGAGACCTTATTCATTCTCAA 2280
Oy 2281 CTCTAGCCCTCAGACACTCCTGTGGCAGGAATGAAAAAGAGCTTCCGTGCAAAAG 2340
Db 2281 CTCTAGCCCTCAGACACTCCTGTGGCAGGAATGAAAAAGAGCTTCCGTGCAAAAG 2340
Oy 2341 GGTACAGCCTTCCACCCCGCCCTCCCTGACCTCTGTCCTCTCCAGTTCATTCT 2400
Db 2341 GGTACAGCCTTCCACCCCGCCCTCCCTGACCTCTGTCCTCTCCAGTTCATTCT 2400
Oy 2401 GGAACCCAGCCAGGAGCAACAGTGGGCGCCCAAGAGGCAAGGATCTCAGGCCCCA 2460
Db 2401 GGAACCCAGCCAGGAGCAACAGTGGGCGCCCAAGAGGCAAGGATCTCAGGCCCCA 2460
Oy 2461 GCGCGGAGGAGCTGGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCACCAGGTCAG 2520
Db 2461 GCGCGGAGGAGCTGGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCACCAGGTCAG 2520
Oy 2521 TCTTTGCTGCTGCTCCCGACACCTCTGTGACACACAGCCAGATCAGAGGACAGGCCA 2580
Db 2521 TCTTTGCTGCTGCTCCCGACACCTCTGTGACACACAGCCAGATCAGAGGACAGGCCA 2580
Oy 2581 GAGATAGCTCTTTTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db 2581 GAGATAGCTCTTTTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Oy 2641 GTATCTGCTTGTACTGAGAAATAAACATTTTCAATAAAAAA 2700
Db 2641 GTATCTGCTTGTACTGAGAAATAAACATTTTCAATAAAAAA 2700
Oy 2701 AAAA 2704
Db 2701 AAAA 2704

RESULT 3
AAD36063
ID AAD36063 standard; cdna; 2618 BP.
XX AAD36063;
AC
XX
DT 09-AUG-2002 (first entry)
XX
DE Murine dual-specificity phosphatase 15 (DSP-15) cdna.
XX
KW Murine: dual-specificity phosphatase 15; DSP15; anti-allergic; cytosolic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; gene; ss.
XX
OS Mus musculus.
XX
FH key Location/Qualifiers
FT CDS 35..1450
FT /tag= a
FT /product= "Murine DSP-15 protein"
XX
PN MO200224740-A2.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001MO-US29406.
XX
PR 19-SEP-2000; 2000US-233833P.
PR 18-SEP-2001; 2001US-0955732.
XX
PA (CEPT-) CEPTYR INC.

XX
PI Luche RM, Wei B:
XX
DR WPI: 2002-394127/42.
XX P-PSDB; AAE22733.
XX
PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
PT diseases
XX
PS
XX
SS
XX
CC The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphoserine/threonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides, may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is murine DSP-15 cdna.
XX
SQ Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other:
XX
Query Match 96.8%; Score 2618; DB 24; Length 2618:
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 60 CCGGTGCCAGCCAGGTGCTCGCGGCTGCTCCATGAGCCCTGTGACAGTGAAGCCGTTTC 119
Db 1 CCGGTGCCAGCCAGGTGCTCGCGGCTGCTCCATGAGCCCTGTGACAGTGAAGCCGTTTC 60
Oy 120 GCGCCCGGGGAGGCGGCTCCAGCGCCGCTGGGGCCCTGGGACCAAGCGGTTCACAGAG 179
Db 61 GCGCCCGGGGAGGCGGCTCCAGCGCCGCTGGGGCCCTGGGACCAAGCGGTTCACAGAG 120
Oy 180 GAGTCGACTCCAGCAAGGAGAGCTTGGGTGCTGCTGGGGCTGCTGGGACTCA 239
Db 121 GAGTCGACTCCAGCAAGGAGAGCTTGGGTGCTGCTGGGGCTGCTGGGACTCA 180
Oy 240 GGATGAGAGGAGCAATGATGACAGCAGAGGCCAGTTCTGAGCCACAGAGAGGCCCC 299
Db 181 GGATGAGAGGAGCAATGATGATGACAGCAGAGGCCAGTTCTGAGCCACAGAGAGGCCCC 240
Oy 300 GAGTGAGAGAGGCTCCAGCGGGGACAGACAGACTTCCGGGCAAGATCCAGATCCCCA 359
Db 241 GAGTGAGAGAGGCTCCAGCGGGGACAGACAGACTTCCGGGCAAGATCCAGATCCCCA 300
Oy 360 GAAGCAGAGAGAGCAGAGAGCAGACCTGACCTCATGTGTACAGCTGCTGAGAGCCGACAGA 419
Db 301 GAAGCAGAGAGAGCAGAGAGCAGACCTGACCTCATGTGTACAGCTGCTGAGAGCCGACAGA 360
Oy 420 TGACATCCGCTGGCAGGCCAGCTGGAGGACCCCGGCTCCCGGGCTCCGCTACCTGCT 479
Db 361 TGACATCCGCTGGCAGGCCAGCTGGAGGACCCCGGCTCCCGGGCTCCGCTACCTGCT 420
Oy 480 GGTAGTTCTACAGAGAGAGAGAGTCTAGCCAGATATAGAGAGGTCTCTGGGGCT 539
Db 421 GGTAGTTCTACAGAGAGAGAGAGTCTAGCCAGATATAGAGAGGTCTCTGGGGCT 480
Oy 540 GGATTTCCCTGACAGAGAGCTCCCGCAGCTGACCTGGGCTGTGCTGCCCCCTGTGGAG 599
Db 481 GGATTTCCCTGACAGAGAGCTCCCGCAGCTGACCTGGGCTGTGCTGCCCCCTGTGGAG 540
Oy 600 TGACACCCAGGTGTACTTAGATGAGAGAGCGGGGCTTACAGCTGACGCTGTGGGCAAG 659
|||||

Dh 541 TGACACCCAGGTGTACTAGATGAGACGGGGGCTTCAAGCTGACGCTGGTGGGCAAG 600
Oy 660 CCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCACTCCAGATATTGACCA 719
Dh 601 CCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCACTCCAGATATTGACCA 660
Oy 720 ACCATGTGAGCAGCTCTAGCAGCGGCTTGTACCGGGTGGCACTGCCCTACCTGGGC 779
Dh 661 ACCATGTGAGCAGCTCTAGCAGCGGCTTGTACCGGGTGGCACTGCCCTACCTGGGC 720
Oy 780 CAGCCACTTCAAGGAGAGCTGAACCGAACAGAGCTCCCTCAATGATGTGAGCGCTAT 839
Dh 721 CAGCCACTTCAAGGAGAGCTGAACCGAACAGAGCTCCCTCAATGATGTGAGCGCTAT 780
Oy 840 GGCAGCAGCTGAGCTGTGCGGCTCCAGCGCCAGCGCTGGCGGCTCTCTAGAAACAGA 899
Dh 781 GGCAGCAGCTGAGCTGTGCGGCTCCAGCGCCAGCGCTGGCGGCTCTCTAGAAACAGA 840
Oy 900 GCAGATGAGAGAGGAGTCCGTGTGAGTGTGAAAGTGTGATGTGATGAGTACCTGGA 959
Dh 841 GCAGATGAGAGAGGAGTCCGTGTGAGTGTGAAAGTGTGATGTGATGAGTACCTGGA 900
Oy 960 GAGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCCA 1019
Dh 901 GAGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCCA 960
Oy 1020 GCAGTACCGTGTGATGACCAACCATGCTGTGTGTGGCAGCAGCGGAGCCAGAC 1079
Dh 961 GCAGTACCGTGTGATGACCAACCATGCTGTGTGTGGCAGCAGCGGAGCCAGAC 1020
Oy 1080 CTCGCCATCTTCCCGACCTTACCTGGGCTCAGAGTGGAAAGCGCAAACTGGAGGA 1139
Dh 1021 CTCGCCATCTTCCCGACCTTACCTGGGCTCAGAGTGGAAAGCGCAAACTGGAGGA 1080
Oy 1140 GCTGAGAGAGAGAGGAGTACCACATCTTGAACATGAGCGCGGAGATTGACAACTTCTA 1199
Dh 1081 GCTGAGAGAGAGAGGAGTACCACATCTTGAACATGAGCGCGGAGATTGACAACTTCTA 1140
Oy 1200 CCTGTAGCGCTTACCTACCAATGTGGGCTCTGGGATGAGAGTCCGCCAGCTGCT 1259
Dh 1141 CCTGTAGCGCTTACCTACCAATGTGGGCTCTGGGATGAGAGTCCGCCAGCTGCT 1200
Oy 1260 GCGGCACTGGAAGAGAGAGCAGCGCTTCAATGAGTGGAGAGCAGAGGAGCCAGCT 1319
Dh 1201 GCGGCACTGGAAGAGAGAGCAGCGCTTCAATGAGTGGAGAGCAGAGGAGCCAGCT 1260
Oy 1320 GCTGTGTCACTGCAAGATGGGCTCAGCGCTCAGCGGACAGTCTGGCTATGCCAT 1379
Dh 1261 GCTGTGTCACTGCAAGATGGGCTCAGCGCTCAGCGGACAGTCTGGCTATGCCAT 1320
Oy 1380 GAAGCAGTACGAATGAGCTGTGAGCAGGCGCTGCGCCACAGTGCAGAGCTCGGCCAT 1439
Dh 1321 GAAGCAGTACGAATGAGCTGTGAGCAGGCGCTGCGCCACAGTGCAGAGCTCGGCCAT 1380
Oy 1440 CCGCGCGCCCAACCTGTGCTTCTGTGGCCAGCTGCAAGATGTACCAAGGCACTCTGACGGC 1499
Dh 1381 CCGCGCGCCCAACCTGTGCTTCTGTGGCCAGCTGCAAGATGTACCAAGGCACTCTGACGGC 1440
Oy 1500 CAGAACCTGAGGCTGTGGGAGAGAGGTTGTAGGCATGGAAGAGACAGCAGGAGCC 1559
Dh 1441 CAGAACCTGAGGCTGTGGGAGAGAGGTTGTAGGCATGGAAGAGACAGCAGGAGCC 1500
Oy 1560 CGAAGAGAGCTGTGGGCAAGGCGCAAGTATAAAGCTCCAGAGGAGTATGAGTCCATCA 1619
Dh 1501 CGAAGAGAGCTGTGGGCAAGGCGCAAGTATAAAGCTCCAGAGGAGTATGAGTCCATCA 1560
Oy 1620 GTCTTGTGAGCCCTCTTGTGAGCTGTGAGAGCAGCTGAGACAGTGAATGCCAGAG 1679
Dh 1561 GTCTTGTGAGCCCTCTTGTGAGCTGTGAGAGCAGCTGAGACAGTGAATGCCAGAG 1620
Oy 1680 TCTTCTTCCAGAGTCTTCAATGAAGAGCTGTGAGAGCCCTTCCAGAGCTTCCAA 1739
Dh 1621 TCTTCTTCCAGAGTCTTCAATGAAGAGCTGTGAGAGCCCTTCCAGAGCTTCCAA 1680

Oy 1740 GGACCAAGGAGGCCAGCAGTGTGACAGAGGGGCTCAAGCTGCTGCTGAGTATGCCAGT 1799
Dh 1681 GGACCAAGGAGGCCAGCAGTGTGACAGAGGGGCTCAAGCTGCTGCTGAGTATGCCAGT 1740
Oy 1800 CAGTGTATACCTCCAGGCGAGTGCCTGTGGTGGCAACCGGAGCCAGGCTTCCAGAGC 1859
Dh 1741 CAGTGTATACCTCCAGGCGAGTGCCTGTGGTGGCAACCGGAGCCAGGCTTCCAGAGC 1800
Oy 1860 AGAGCAGAGGCGAGGCGAGGCGAGAGAGCCCTGATTTCTCTACGCCAGGTTCC 1919
Dh 1801 AGAGCAGAGGCGAGGCGAGGCGAGAGAGCCCTGATTTCTCTACGCCAGGTTCC 1860
Oy 1920 GGAAGTGTGTAGACAGGCCAGCGTGCATGACAGTGGAGAGAGAGGCGAGGCTGAGCC 1979
Dh 1861 GGAAGTGTGTAGACAGGCCAGCGTGCATGACAGTGGAGAGAGGCGAGGCTGAGCC 1920
Oy 1980 TCACACATGCGCAGCTCCCTGACACTGAGAGAGATCCAAAGCTCTTGGAGAAACAC 2039
Dh 1921 TCACACATGCGCAGCTCCCTGACACTGAGAGAGATCCAAAGCTCTTGGAGAAACAC 1980
Oy 2040 CTCACGTGTGTGCGCAGACATTCCTCTCAGCTCCGCCCATACCGCTCAGTACAGCT 2099
Dh 1981 CTCACGTGTGTGCGCAGACATTCCTCTCAGCTCCGCCCATACCGCTCAGTACAGCT 2040
Oy 2100 CACCTCCCAACCCCTGTCACTACAGGCTCAGCTCCAGCCCTGTCTACAGCTCAGCTC 2159
Dh 2041 CACCTCCCAACCCCTGTCACTACAGGCTCAGCTCCAGCCCTGTCTACAGCTCAGCTC 2100
Oy 2160 CTCACGCTTAAGTCCAGAGGCCATGTGCTGTCTGCAAGGAGCTCAAGACTTCTAAGT 2219
Dh 2101 CTCACGCTTAAGTCCAGAGGCCATGTGCTGTCTGCAAGGAGCTCAAGACTTCTAAGT 2160
Oy 2220 GGATGTGTAGAGGAGCAGTGAAGTACCTTTGGGGCAACAGCACCCTAGTTCACTTCA 2279
Dh 2161 GGATGTGTAGAGGAGCAGTGAAGTACCTTTGGGGCAACAGCACCCTAGTTCACTTCA 2220
Oy 2280 ACTTACGCCCTCAGACATCACCTGTGGCAGGAGTGAAGAAAGAGACTTCCGTGCAAAA 2339
Dh 2221 ACTTACGCCCTCAGACATCACCTGTGGCAGGAGTGAAGAAAGAGACTTCCGTGCAAAA 2280
Oy 2340 GGGTCAAGCTCCCAACCCCGCCCTCTCTGACACTGCTGTCTCTCCAGTTCAATTC 2399
Dh 2281 GGGTCAAGCTCCCAACCCCGCCCTCTCTGACACTGCTGTCTCTCCAGTTCAATTC 2340
Oy 2400 TGAACCAAGCAGGCCAGGCAACAGTGGCCCAAGGCAAGGAGAGTCTCAGGCCCC 2459
Dh 2341 TGAACCAAGCAGGCCAGGCAACAGTGGCCCAAGGCAAGGAGAGTCTCAGGCCCC 2400
Oy 2460 AGCGGAGGAGCTGGAAGGCTGGCAGATGGCTTCCATCCAGCTCCAGCGGTCCAG 2519
Dh 2401 AGCGGAGGAGCTGGAAGGCTGGCAGATGGCTTCCATCCAGCTCCAGCGGTCCAG 2460
Oy 2520 GTCTTGTGTGCTGTCCCGCAGACCTCTGTGACACAGCAGCAGATCAGAGGAGCCAGG 2579
Dh 2461 GTCTTGTGTGCTGTCCCGCAGACCTCTGTGACACAGCAGCAGATCAGAGGAGCCAGG 2520
Oy 2580 AGAGATAGTCTTCTTTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2639
Dh 2521 AGAGATAGTCTTCTTTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2580
Oy 2640 AGTATGTGCTTGTACTGAGAAATTAACACATTTTC 2677
Dh 2581 AGTATGTGCTTGTACTGAGAAATTAACACATTTTC 2618

RESULT 4
AAF30485
ID AAF30485 standard; cDNA: 2781 BP.
XX
XX AAF30485;
AC
XX 29-MAY-2001 (first entry)

XX Human protein phosphatase and kinase protein-10 cDNA 5039718CBI.
 XX DE
 KW Protein phosphatase and kinase protein; PPHK-10; human;
 KW gastrointestinal disorder; immune system disorder;
 KW neurological disorder; cell proliferative disorder; cancer;
 KW diagnosis; therapy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 87..2066
 FT sig_peptide /tag= a
 FT 87..143
 FT mat_peptide /tag= b
 FT 144..2063
 FT misc_feature /tag= c
 FT 219..279
 FT /tag= d
 FT /note= "unique fragment"
 FT misc_feature 921..980
 FT /tag= e
 FT /note= "unique fragment"
 XX
 PN W0200120004-A2.
 XX
 PD 22-MAR-2001.
 XX
 PF 14-SEP-2000; 2000WO-US25515.
 XX
 PR 15-SEP-1999; 99US-0154141.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAI;
 XX
 DR WPI: 2001-244811/25.
 DR P-PSDB: AAB20331.
 XX
 PT Novel human protein phosphatase and kinase proteins for diagnosis,
 PT treatment and prevention of gastrointestinal, immune system,
 PT neurological and cell proliferative disorders -
 XX
 PS Claim 5; Page 101-102; 103pp; English.
 XX
 CC The present sequence is that of cDNA encoding novel human
 CC protein phosphatase and kinase protein PPHK-10 (see AAB20331).
 CC The cDNA was initially identified in Incyte Clone ID No. 5039718CBI,
 CC from a colon tumour tissue library. Tissues that express PPHK-10
 CC (as a fraction of total tissues expressing PPHK-10) include
 CC reproductive (0.343), gastrointestinal (0.194) and haemopoietic
 CC or immune (0.134). Diseases or conditions associated with tissues
 CC expressing PPHK-10 (as a fraction of total tissues expressing
 CC PPHK-10) include cancer (0.552), inflammation or trauma (0.314) or
 CC cell proliferation (0.090). The encoded protein shows homology to
 CC Drosophila melanogaster MAP kinase phosphatase. The invention
 CC provides human PPHK-1 to -11 polypeptides (see AAB20322-32) and
 CC polynucleotides (see AAF30476-86). It also provides expression
 CC vectors, host cells, antibodies, agonists and antagonists, as well
 CC as methods for diagnosing, treating or preventing disorders
 CC associated with expression of PPHK, including gastrointestinal
 CC disorders, immune system disorders, neurological disorders and cell
 CC proliferative disorders, including cancer.
 XX
 SQ Sequence 2781 BP: 576 A; 906 C; 820 G; 479 T; 0 other;
 Query Match 94.8%; Score 2562.8; DB 22; Length 2781;
 Best Local Similarity 96.4%; Pred. No. 0;
 Matches 2672; Conservative 0; Mismatches 2; Indels 98; Gaps 1;
 OY 17 GCGGGTCCAGAGACTGTCCGCGGGTTGAGGAGAGGGCCGTGCCCGTCCAGCCAGCT 76
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 10 GCGGGTCCAGAGACTGTCCGCGGGTTGAGGAGAGGGCCGTGCCCGTCCAGCCAGCT 69
 QY 77 GCTGCGGGCTGGCTCCATATGGCCCGTGCACAGTGAAGCCGTTCGGCCCGGAGCGGCG 136
 Db 70 GCTGCGGGCTGGCTCCATATGGCCCGTGCACAGTGAAGCCGTTCGGCCCGGAGCGGCG 129
 QY 137 CCTCCACGCCCGTGGGGCCCTGGGACACAGGCGGCTCCAGCGAAGAGTGCATCCAGCGAA 196
 Db 130 CCTCCACGCCCGTGGGGCCCTGGGACACAGGCGGCTCCAGCGAAGAGTGCATCCAGCGAA 189
 QY 197 GGCAGAGCTTTGGGGTGTCTCCGTGGGGCTGTCTGGGACTGCAGAGTGAAGGGACAAATG 256
 Db 190 GGCAGAGCTTTGGGGTGTCTCCGTGGGGCTGTCTGGGACTGCAGAGTGAAGGGACAAATG 249
 QY 257 ATGATGACAGAGAGGCGGCTGTGAGCCAAAGAGAGAGGCGGCGGATGAGAGAGCTCC 316
 Db 250 ATGATGACAGAGAGGCGGCTGTGAGCCAAAGAGAGAGGCGGCGGATGAGAGAGCTCC 309
 QY 317 ACGGGACACAGAGAGCTTCGGGCAAGGATCCAGAGTCCCGAGAGGAGAGAGAGAGAGAG 376
 Db 310 ACGGGACACAGAGAGCTTCGGGCAAGGATCCAGAGTCCCGAGAGGAGAGAGAGAGAGAG 369
 QY 377 GGCAGACCTTCACCTCATGTGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAG 436
 Db 370 GGCAGACCTTCACCTCATGTGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAG 429
 QY 437 CCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAG 496
 Db 430 CCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAG 489
 QY 497 AAGGAGAGCTCTGAGAGCAGAGATGAGAGCTGCTCCGCGGCTGGATTTCCCTGACAGA 556
 Db 490 AAGGAGAGCTCTGAGAGCAGAGATGAGAGCTGCTCCGCGGCTGGATTTCCCTGACAGA 549
 QY 557 GCTCCCGCAGCTGCACCCCTGGGCTGTGCTTGCCTCTGAGATGACAGCCAGGTGTACT 616
 Db 550 GCTCCCGCAGCTGCACCCCTGGGCTGTGCTTGCCTCTGAGATGACAGCCAGGTGTACT 609
 QY 617 TAGATGAGAGAGGGGGCTTCAGCGGTGAGCTGTGGGGCAAGCGGATCTTCAAGCCCA 676
 Db 610 TAGATGAGAGAGGGGGCTTCAGCGGTGAGCTGTGGGGCAAGCGGATCTTCAAGCCCA 669
 QY 677 TCTCCATCCAGACCATGTGGGCGCACACTCCAGATATTCACCAAGCATGTAGGACAGCTC 736
 Db 670 TCTCCATCCAGACCATGTGGGCGCACACTCCAGATATTCACCAAGCATGTAGGACAGCTC 729
 QY 737 TAGGACGCGGCTGTGTACCGGGGTGGAGTGCCTTCACCTGGGCCACGCACATACAGAGA 796
 Db 730 TAGGACGCGGCTGTGTACCGGGGTGGAGTGCCTTCACCTGGGCCACGCACATACAGAGA 789
 QY 797 GACTGAACCTCGAACAGAGAGCTCCCAATGATGAGTGAAGCGGTATGCGGACCTGAGTCTC 856
 Db 790 GACTGAACCTCGAACAGAGAGCTCCCAATGATGAGTGAAGCGGTATGCGGACCTGAGTCTC 849
 QY 857 TGGCGGCTCCAGAGCGGCGAGGCTGGGGGTCTCTCAGAACAGAGACAGATGAGAGGCGGA 916
 Db 850 TGGCGGCTCCAGAGCGGCGAGGCTGGGGGTCTCTCAGAACAGAGACAGATGAGAGGCGGA 909
 QY 917 TCCGTGCTGAGCTGTGAAGAGTGTGATGTCAAGTGAAGCTGAGAGAGTGCATCTTCCAAG 976
 Db 910 TCCGTGCTGAGCTGTGAAGAGTGTGATGTCAAGTGAAGCTGAGAGAGTGCATCTTCCAAG 969
 QY 977 AGATTCGCGCAGAGCTGTGAGAGTGTGGGCTGGGGGTCTCCCTCCAGACAGTACCTGACTTCA 1036
 Db 970 AGATTCGCGCAGAGCTGTGAGAGTGTGGGCTGGGGGTCTCCCTCCAGACAGTACCTGACTTCA 1029
 QY 1037 TCGACAAACCAAGATCTGCTGTGTGAGACAGCGGAGCGGACCTCCCGCATCTTCCGCC 1096
 Db 1030 TCGACAAACCAAGATCTGCTGTGTGAGACAGCGGAGCGGACCTCCCGCATCTTCCGCC 1089
 QY 1097 ACCTTACCTGGGCTCAGAGTGAAGCAGCAGAACTTGAAGAGCTGCAGAGAGAGAGG 1156
 Db 1090 ACCTTACCTGGGCTCAGAGTGAAGCAGCAGAACTTGAAGAGCTGCAGAGAGAGAGG 1149

QY 1157 TCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCT 1216
 |||||
 Db 1150 TCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCT 1209
 QY 1217 ACCCAATATGCGCTCTGGGATGAGAGATCGGCCAGTGTCTGCCGACCTGGAAGGAGA 1276
 |||||
 Db 1210 ACCCAATATGCGCTCTGGGATGAGAGATCGGCCAGTGTCTGCCGACCTGGAAGGAGA 1269
 QY 1277 GCGACCGCTTCTTGAAGTGTGCAAGAGCAGAGGCAACCAGCTGTCTGTCACCTGCAAGA 1336
 |||||
 Db 1270 GCGACCGCTTCTTGAAGTGTGCAAGAGCAGAGGCAACCAGCTGTCTGTCACCTGCAAGA 1329
 QY 1337 TGGGCGTTCAGCCGCTCAGCGGCCACAGTGTGCTATGSCATGAAGAGTAACTGCA 1396
 |||||
 Db 1330 TGGGCGTTCAGCCGCTCAGCGGCCACAGTGTGCTATGSCATGAAGAGTAACTGCA 1389
 QY 1397 GCGTGAAGAGGCGCTTGGCCGACGTGACAGAGCTCCGGGCGCATCGCCGCCCAACCCCTG 1456
 |||||
 Db 1390 GCGTGAAGAGGCGCTTGGCCGACGTGACAGAGCTCCGGGCGCATCGCCGCCCAACCCCTG 1449
 QY 1457 GCTTCTGGCGCAGCTGCAATCTACAGGGCATCTGACG----- 1497
 |||||
 Db 1450 GCTTCTGGCGCAGCTGCAATCTACAGGGCATCTGACG----- 1509
 QY 1498 ----- 1497
 Db 1510 TCTGGAGAGAGAAATGGGTGGGTCTCCCAAGAGGACACCAGCCCTGAATCTCTA 1569
 QY 1498 -----GCCAGACCTGAGGGTGGTGGGGAGGAGGTTGTAGGA 1538
 |||||
 Db 1570 CACCAATTCACCTCTTCCGCGCAGAACTGAGGGTGTGGGGAGAAAGTTGTAGGA 1629
 QY 1539 TGGAAAGAGAGCCAGGAGGCGCCGAAGAGAGCTGGGCCACGCGCCACGATTAACCTCC 1598
 |||||
 Db 1630 TGGAAAGAGAGCCAGGAGGCGCCGAAGAGAGCTGGGCCACGCGCCACGATTAACCTCC 1689
 QY 1599 GAGGGGTATGAGGTCATCATGATCTTCTGAGAGCCCTTCTGAGCTGAGAGACCTCAG 1658
 |||||
 Db 1690 GAGGGGTATGAGGTCATCATGATCTTCTGAGAGCCCTTCTGAGCTGAGAGACCTCAG 1749
 QY 1659 AGACCAAGTATGAGGTCAGAGAGGCTTCTCCAGAGAGTTCACATGAAAGACCTCTGC 1718
 |||||
 Db 1750 AGACCAAGTATGAGGTCAGAGAGGCTTCTCCAGAGAGTTCACATGAAAGACCTCTGC 1809
 QY 1719 AGCCCTTCCACAGCTTGAAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1778
 |||||
 Db 1810 AGCCCTTCCACAGCTTGAAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1869
 QY 1779 CTGCGCTGAAGTCCCGCAGTCACTGTTACCTCCAGGCACTGCCGTGTGGCCAAAC 1838
 |||||
 Db 1870 CTGCGCTGAAGTCCCGCAGTCACTGTTACCTCCAGGCACTGCCGTGTGGCCAAAC 1929
 QY 1839 GGAACCCAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1898
 |||||
 Db 1930 GGAACCCAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1989
 QY 1899 TTTCTCTTACGCGCAGTGTCCGGAAGTGTGAGAGCAGGCGAGGCTGATGACAGTGGAG 1958
 |||||
 Db 1990 TTTCTCTTACGCGCAGTGTCCGGAAGTGTGAGAGCAGGCGAGGCTGATGACAGTGGAG 2049
 QY 1959 AGGAGGCGAGGCTGAGCCCTCACACATGCCACAGCTCCCTGACACTGAAGAGATCC 2018
 |||||
 Db 2050 AGGAGGCGAGGCTGAGCCCTCACACATGCCACAGCTCCCTGACACTGAAGAGATCC 2109
 QY 2019 ACAACTCTTGGAGAAACACACCTCAGCTGTGCGGACACACATCTCTCAGTCCGCC 2078
 |||||
 Db 2110 ACAACTCTTGGAGAAACACACCTCAGCTGTGCGGACACACATCTCTCAGTCCGCC 2169
 QY 2079 CCATACCCTGCTACTAGAGCTCAGCTCCACACCTGTCTACTAGAGGCTCAGCTCCACCC 2138
 |||||
 Db 2170 CCATACCCTGCTACTAGAGCTCAGCTCCACACCTGTCTACTAGAGGCTCAGCTCCACCC 2229

QY 2139 CTGTCTACAGGCTCAGCTCAGCTTACAGCTTAACTGCCAGGCGCCATGTCCTTCCAA 2198
 |||||
 Db 2230 CTGTCTACAGGCTCAGCTCAGCTTACAGCTTAACTGCCAGGCGCCATGTCCTTCCAA 2289
 QY 2199 GGGCTCAGAGCTTCTTAACTGGATGTGGTAGAGAGGACTGAAGTACTTGTGGGGCAAC 2258
 |||||
 Db 2290 GGGCTCAGAGCTTCTTAACTGGATGTGGTAGAGAGGACTGAAGTACTTGTGGGGCAAC 2349
 QY 2259 AGCACCTTAGTTTCTTCTTAACTGTAGCCCTGCACACTCAGCTGTGGCAGGAAATGAA 2318
 |||||
 Db 2350 AGCACCTTAGTTTCTTCTTAACTGTAGCCCTGCACACTCAGCTGTGGCAGGAAATGAA 2409
 QY 2319 ACAGAGCTTCCCGTCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2378
 |||||
 Db 2410 ACAGAGCTTCCCGTCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2469
 QY 2379 TGTCTCTCCAGTTCAATCTCTGAAACAGCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGG 2438
 |||||
 Db 2470 TGTCTCTCCAGTTCAATCTCTGAAACAGCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGG 2529
 QY 2439 CAGGAGGATCTCTAGGCGCCAGCGCGGAGGCTGGAAGGCTGGCAGATCGCTTCCCT 2498
 |||||
 Db 2530 CAGGAGGATCTCTAGGCGCCAGCGCGGAGGCTGGAAGGCTGGCAGATCGCTTCCCT 2589
 QY 2499 CATCCACCTCCAGCGGCTCAGGCTTGTGCTGTCGCCAGACCTCTGTGACACACAGC 2558
 |||||
 Db 2590 CATCCACCTCCAGCGGCTCAGGCTTGTGCTGTCGCCAGACCTCTGTGACACACAGC 2649
 QY 2559 CAGATCACAGGCGCAGGCGCAGAGATAGTCTTCTTGTCTCTTCTGCGCTTGCGCTA 2618
 |||||
 Db 2650 CAGATCACAGGCGCAGGCGCAGAGATAGTCTTCTTGTCTCTTCTGCGCTTGCGCTA 2709
 QY 2619 GTTCAGTTTCTTATAGCTTACAGTATCTGGCTTGTACTGAGAAATTAACACATTTTCA 2678
 |||||
 Db 2710 GTTCAGTTTCTTATAGCTTACAGTATCTGGCTTGTACTGAGAAATTAACACATTTTCA 2769
 QY 2679 TAAAAAATAA 2690
 |||||
 Db 2770 TAAAAAATAA 2781
 |||||
 RESULT 5
 AB073250
 ID AB073250 standard: cDNA; 2852 BP.
 XX AC AB073250;
 XX DT 30-SEP-2002 (first entry)
 XX DE Human MAP kinase phosphatase splice form 2 cDNA sequence SEQ ID NO:2.
 XX KW Human; phosphatase; mitogen activated protein kinase phosphatase;
 XX KW MAP kinase; enzyme; chromosome 11; gene; ss.
 XX OS Homo sapiens.
 XX FH key Location/Qualifiers
 XX FT 5'UTR 1..56
 XX FT CDS /*tag= a
 XX FT /*tag= b
 XX FT /*tag= c
 XX FT 3'UTR
 XX PN MO200242436-A2.
 XX PD 30-MAY-2002.
 XX PF 07-NOV-2001; 2001WO-US42995.
 XX PR 20-NOV-2000; 2000US-0715177.
 PR 18-JAN-2001; 2001US-0761640.

[illegible]

DB	2762	AAAA	2765
RESULT	6		
AAD36061			
ID	AAD36061	standard; cDNA;	2718 BP.
AC			
XX	AAD36061;		
DT	09-AUG-2002	(first entry)	
DE			
XX			
KM	Human dual-specificity phosphatase 15 (DSP-15) cDNA.		
KW	Human; dual-specificity phosphatase 15; DSP-15; antiallergic; cytosolic;		
KW	immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;		
KW	signal transduction; cell proliferation; Duchenne muscular dystrophy;		
KW	cell cycle abnormality; graft-versus-host disease; autoimmune disease;		
KW	metabolic disease; allergy; screening; chromosome 11q; gene; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	35..2014	
FT		/*tag= a	
FT		/product= "Human DSP-15 protein"	
PN	WO200224740-A2.		
PD			
XX	28-MAR-2002.		
PF	19-SEP-2001; 2001WO-US29406.		
XX			
PR	19-SEP-2000; 2000US-233833P.		
PR	18-SEP-2001; 2001US-095573Z.		
XX			
PA	(CEPT-) CEPTYR INC.		
XX			
PI	Luche RM, Wei B:		
DR	WPI: 2002-394127/42.		
P-PSDB:	AAE22729.		
PT	New dual-specificity phosphatase 15 polypeptide and polynucleotides,		
PT	useful for treating e.g. Duchenne muscular dystrophy, cancer,		
PT	graft-versus-host disease, autoimmune diseases, allergies, metabolic		
PT	diseases -		
PS	Claim 7; Fig 1; 91pp; English.		
XX			
CC	The invention relates to a new isolated dual-specificity phosphatase 15		
CC	(DSP-15) polypeptide which retains the ability to dephosphorylate an		
CC	activated MAP (mitogen activated protein) kinase. DSPs are phosphatases		
CC	that dephosphorylate both phosphotyrosine and phosphothreonine/serine		
CC	residues. DSP-15 polypeptides may be used to identify agents that		
CC	modulate DSP-15 activity, where such agents may inhibit or enhance signal		
CC	transduction via a MAP-kinase cascade, leading to cell proliferation. DSP		
CC	polypeptides, modulating agents, and/or polynucleotides encoding the		
CC	polypeptides may be used to modulate DSP-15 activity in a patient, and to		
CC	ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-		
CC	versus-host disease, autoimmune diseases, allergies, metabolic diseases,		
CC	abnormal cell growth, abnormal cell proliferation and cell cycle		
CC	abnormalities. DSP-15 alternate form polypeptides are useful in screening		
CC	assays for modulators of enzyme activity and/or substrate binding. The		
CC	present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on		
CC	chromosome 11q.		
XX			
SQ	Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;		
Best Match	92.9%; Score 2512; DB 24; Length 2718;		
Query Similarity	96.4%; Pred. No. 0;		
Matches 2620; Conservative	0; Mismatches 0; Indels 98; Gaps 1		
CGGTCGCCAGCCCAAGTCTCAGCGGCCTGGCTCATATGCCCCTTGTCACAGTAGCCGTTTC	119		

|||||
Db 1 CCGGTCACGACCAGGTGCTCGGGCTGCTCCATGGCCCTGCTCACAGTAGCCGTTTC 60
OY 120 GCCCGCGGACAGCGGCGCTCCACGCGCGTGGGGCCCTGGGACACAGGGGTCACAGGAAG 179
Db 61 GCGCCCGGACAGCGGCGCTCCACGCGCGTGGGGCCCTGGGACACAGGGGTCACAGGAAG 120
OY 180 GAGTCGACTCCAGCAGAGCAGAGCGCTTGGGGTCTCCCTGGGGGCTGTTCTGGACTCA 239
Db 121 GAGTCGACTCCAGCAGAGCAGAGCGCTTGGGGTCTCCCTGGGGGCTGTTCTGGACTCA 180
OY 240 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 299
Db 181 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 240
OY 300 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 359
Db 241 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 300
OY 360 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 419
Db 301 GAGTCGAGGGGCAATGATGATGACAGAGGCGCATGTTCTGAGCCACAGAGAGAGCCCC 360
OY 420 TCACATCCGCTCGCAGCCGCTGAGGACACCCGCGCTCCCGGCTCCGCTACCTGCT 479
Db 361 TCACATCCGCTCGCAGCCGCTGAGGACACCCGCGCTCCCGGCTCCGCTACCTGCT 420
OY 480 GGTAGTTCTACAGGAG 539
Db 421 GGTAGTTCTACAGGAG 480
OY 540 GGTAGTTCTACAGGAG 599
Db 481 GGTAGTTCTACAGGAG 540
OY 600 TCACATCCGCTCGCAGCCGCTGAGGACACCCGCGCTCCCGGCTCCGCTACCTGCT 659
Db 541 TCACATCCGCTCGCAGCCGCTGAGGACACCCGCGCTCCCGGCTCCCGGCTCCGCTACCTGCT 600
OY 660 CCGGATCTTCAAGCCCATTCATCCAGACAGCATGTTGGGCGCACATCCAGATTTGACCA 719
Db 601 CCGGATCTTCAAGCCCATTCATCCAGACAGCATGTTGGGCGCACATCCAGATTTGACCA 660
OY 720 AGCATGTGAGGAGCTCTAGGAGGCGGCTTGTACCGGGTGGCAGTGCCTACCTGAGGC 779
Db 661 AGCATGTGAGGAGCTCTAGGAGGCGGCTTGTACCGGGTGGCAGTGCCTACCTGAGGC 720
OY 780 CAGCCACTACACAG 839
Db 721 CAGCCACTACACAG 780
OY 840 GCGGACCTGAGGAGCTCTGCGGGGCTCCCAAGGCGGAGAGAGAGAGAGAGAGAGAG 899
Db 781 GCGGACCTGAGGAGCTCTGCGGGGCTCCCAAGGCGGAGAGAGAGAGAGAGAGAGAG 840
OY 900 GCAGATGAG 959
Db 841 GCAGATGAG 900
OY 960 GAGTGTCACTTCAAAAGAGATCCGCCAGGCTCTGAGAGCTGCGCTGGGGCTCCCTCCA 1019
Db 901 GAGTGTCACTTCAAAAGAGATCCGCCAGGCTCTGAGAGCTGCGCTGGGGCTCCCTCCA 960
OY 1020 GAGTGTCACTTCAAAAGAGATCCGCCAGGCTCTGAGAGCTGCGCTGGGGCTCCCTCCA 1079
Db 961 GAGTGTCACTTCAAAAGAGATCCGCCAGGCTCTGAGAGCTGCGCTGGGGCTCCCTCCA 1020
OY 1080 CTCGCGCATCTTCCCGCAGCTCTACCTGGGCTCAGAGTGAAGAGAGAGAGAGAGAGAG 1139
Db 1021 CTCGCGCATCTTCCCGCAGCTCTACCTGGGCTCAGAGTGAAGAGAGAGAGAGAGAG 1080
OY 1140 GCTGAG 1199
|||||

Db 1081 GCTGAG 1140
OY 1200 CCTGAGAGCTTACCTACCTACCAATGTTGGCCCTGGGAGTGAAGAGTGGCCAGCTGCT 1259
Db 1141 CCTGAGAGCTTACCTACCTACCAATGTTGGCCCTGGGAGTGAAGAGTGGCCAGCTGCT 1200
OY 1260 GCGGACTGGAAG 1319
Db 1201 GCGGACTGGAAG 1260
OY 1320 GCTGCTCACTCAAGAGTGGGCTCAGCCGCTCAGGCGCCACAGTGTGCTTATGCTCAT 1379
Db 1261 GCTGCTCACTCAAGAGTGGGCTCAGCCGCTCAGGCGCCACAGTGTGCTTATGCTCAT 1320
OY 1380 GAGCAGTACGATGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439
Db 1321 GAGCAGTACGATGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
OY 1440 GCGCGGCCCCAACCCCTGGCTTCTGCGCCAGCTGAGATCTACAGAGGAGATCTGACG -- 1497
Db 1381 GCGCGGCCCCAACCCCTGGCTTCTGCGCCAGCTGAGATCTACAGAGGAGATCTGACGCGC 1440
OY 1498 ----- 1497
Db 1441 CAGCGCCAGAGCCATGTCGTGGAGCAGAAAGTGGGTGGGTCTCCACAGAGAGAGAGAG 1500
OY 1498 ----- 1521
Db 1501 AGCCCTGAGAGTCTACACCATTTCCACCTCTTCCGCGCAGAACTGAGGGTGTGGGA 1560
OY 1522 GAGAGAGAGTGTAGGAGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1581
Db 1561 GAGAGAGAGTGTAGGAGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
OY 1582 GCGAGTATTAAGCTCCGAGGGGCTCATGAGTGCATCACTCTTCTGGAGCCTCTTGA 1641
Db 1621 GCGAGTATTAAGCTCCGAGGGGCTCATGAGTGCATCACTCTTCTGGAGCCTCTTGA 1680
OY 1642 GCTGAGAGACACTCAG 1701
Db 1681 GCTGAGAGACACTCAG 1740
OY 1702 ACATGAAGAGCTCTGACAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAG 1761
Db 1741 ACATGAAGAGCTCTGACAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAG 1800
OY 1762 GAGCAGAGGGGCTCAGAGCTGCTGAGTCCGCGCAGTACAGTGTGTTACCTCCAGGGCAG 1821
Db 1801 GAGCAGAGGGGCTCAGAGCTGCTGAGTCCGCGCAGTACAGTGTGTTACCTCCAGGGCAG 1860
OY 1822 TGCCTGTGAGGAG 1881
Db 1861 TGCCTGTGAGGAG 1920
OY 1882 GCAGGAG 1941
Db 1921 GCAGGAG 1980
OY 1942 GCTGATGAG 2001
Db 1981 GCTGATGAG 2040
OY 2002 GACACTGAAG 2061
Db 2041 GACACTGAAG 2100
OY 2062 TTCTCTCAGCTCGGAG 2121
Db 2101 TTCTCTCAGCTCGGAG 2160
OY 2122 GGCCTACCTCCAGCCCTGTGCTACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2181
Db 2161 GGCCTACCTCCAGCCCTGTGCTACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
|||||

OY	667	TTTAAAGCCCATCTTCATCCAGACCAATGTGGGCCAACACTCCAGTATTGCAACCAAGCATGT	726
Db	661	TTTAAAGCCCATCTTCATCCAGACCAATGTGGGCCAACACTCCAGTATTGCAACCAAGCATGT	688
OY	727	GAGGCAGCTCAGGCACGGCCCTTTGATACGGGGTGGCAGTGGCCCTCACCCTGGGGCAGGCCAC	786
Db	689	-----	688
OY	787	TACCAAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGATGGAGCGCTATGGCCGAC	846
Db	689	-----	688
OY	847	CTGAGATCTCTGGGGCCTCCACAGCCCGAGACCTGGGGGTCTCTCAGAAACAGAGCAATG	906
Db	689	-----GCTCCCTCAGAAACAGAGCAATG	711
OY	907	GAGCAGCGATCCGCTGTGATGTGTGAAGATTGTGAATGTATGAGACTGGAGAGTTC	966
Db	712	GAGCAGCGATCCGCTGTGATGTGTGAAGATTGTGAATGTATGAGACTGGAGAGTTC	771
OY	967	ACTTCCAAAGAGATCGGCCAGGCTCTGGAGCTGGCTGGGGCTCCCTCCAGCATAC	1026
Db	772	ACTTCCAAAGAGATCCGCCAGGCTCTGGAGCTGGGGCTCCCTCCAGCATAC	831
OY	1027	CGTACATTCATCGAACACGATGCTGCTGGTGGTGGACACGGGAGACCGAGCCCTCCGC	1086
Db	832	CGTACATTCATCGAACACGATGCTGCTGGTGGTGGACACGGGAGACCGAGCCCTCCGC	891
OY	1087	ATCTTCCGCCACCTTATACCTGGGCTCAGAGTGAACGCAGCAAACTTGSAGAGCTGCAG	1146
Db	892	ATCTTCCGCCACCTTATACCTGGGCTCAGAGTGAACGCAGCAAACTTGSAGAGCTGCAG	951
OY	1147	AGGAACAGGGTCAACCCACATCTTGAACATGTGGCCGGGAATTTGAACTTCTACCTTAG	1206
Db	952	AGGAACAGGGTCAACCCACATCTTGAACATGTGGCCGGGAATTTGAACTTCTACCTTAG	1011
OY	1207	CGCTTACACTACACATGTGTGCCTCTGTGGATGAGAGATCGGCCACCTGCTGCAGC	1266
Db	1012	CGCTTACACTACACATGTGTGCCTCTGTGGATGAGAGATCGGCCACCTGCTGCAGC	1071
OY	1267	TGGAAGAGAGCACCGCTTCAATTGAGGCTGCAGAGCACAGGACCACCGAGCTGCTGTC	1326
Db	1072	TGGAAGAGAGAGCACCGCTTCAATTGAGGCTGCAGAGCACAGGACCACCGAGCTGCTGTC	1131
OY	1327	CACGTCAAGATGGGCGTACACCGCTCAGAGGGGCACAGTCTGTGGCTATGGCATGAAGAG	1386
Db	1132	CACGTCAAGATGGGCGTACACCGCTCAGAGGGGCACAGTCTGTGGCTATGGCATGAAGAG	1191
OY	1387	TAGCAATGCAAGCTTGAGAGAGGCCCTTGCGCCACGTGCAGAGACTCGGCCCATCGCCGC	1446
Db	1192	TAGCAATGCAAGCTTGAGAGAGGCCCTTGCGCCACGTGCAGAGACTCGGCCCATCGCCGC	1251
OY	1447	CCCAACCCCTGGCTTCCGTCGGCCCAAGTCAGATCTACAGAGGATCTCAAGGCCAGAAC	1506
Db	1252	CCCAACCCCTGGCTTCCGTCGGCCCAAGTCAGATCTCAAGAGGATCTCAAGGCCAGAAC	1311
OY	1507	TGAGGATGTGGGAGAGAGAGTTGTAGGCATGGAAGAGACCAGGAGCCCGGAAGA	1566
Db	1312	TGAGGATGTGGGAGAGAGAGTTGTAGGCATGGAAGAGACCAGGAGCCCGGAAGA	1371
OY	1567	AGAACCTGGG--CCAGGGCCACGTTAATAACCTCCAGGGGTATAGTCCATCAGTCTT	1624
Db	1372	AGAACCTGGGGCCACGGGGCAGTATAAACCTCCAGGGGTATAGTCCATCAGTCTT	1431
OY	1625	CTGAGAGCCCTCTT--GGAGCTGTGAGAGCACCTCAG--AGACAGTACATGCGACAGGCTCT	1682
Db	1432	CTGAGAGCCCTCTTGGGAGGCTGTGAGAGCACCTCAGTATGACACGTAATGCGACAGGCTCT	1491
OY	1683	TCTCTTCCACAGATCTTCACTGAAGAGCTCTGCAGCCCTTCCACAGCTTGCAGGA	1742
Db	1492	TCTCTTCCACAGATCTTCACTGAAGAGCTCTGCAGCCCTTCCACAGCTTGCAGGA	1551

QY	1743	CCAAGGAGGACCAACAGGTGGACAGGGGGCTTAGCTTCCCTGCTGAAGTCCGGCCAGTAG	1802
Db	1552	CCAAGGAGGGCCACAGGTGGACAGGGGGCTTAGCTTCCCTGCTGAAGTCCGGCCAGTAG	1611
QY	1803	TGGTATCCCTCCAGGGGACAGTGGGTGGCCAAACGGACCCAGGCTTCCAGGAGCAGG	1862
Db	1612	TGGTATCCCTCCAGGGGACAGTGGGTGGCCAAACGGACCCAGGCTTCCAGGAGCAGG	1671
QY	1863	AGCAGGGGACAGGGGACAGGGGACAGGGAGCCCGCATTTCTCTAAGCCAGATTCGCGA	1922
Db	1672	AGCAGGGGACAGGGGACAGGGGACAGGGAGCCCGCATTTCTCTAAGCCAGATTCGCGA	1731
QY	1923	AGGTGTGTAGACAGGCCAGCTGCATGTACAGTGTAGAGGAGGGCGAGGCTGTAGCCCTCA	1982
Db	1732	AGGTGTGTAGACAGGCCAGCTGCATGTACAGTGTAGAGGAGGGCGAGGCTGTAGCCCTCA	1791
QY	1983	CACATGCCACAGCTCCCTCTACACTGAAGAAGATCCACAACTCTTTGGAGAAACACCTC	2042
Db	1792	CACATGCCACAGCTCCCTCTACACTGAAGAAGATCCACAACTCTTTGGAGAAACACCTC	1851
QY	2043	ACGTCGTGTGGCCGACACATTTCCCTCAGTCCGCGCCCATACACCGCTACAGCTCTCAC	2102
Db	1852	ACGTCGTGTGGCCGACACATTTCCCTCAGTCCGCGCCCATACACCGCTACAGCTCTCAC	1911
QY	2103	CTCCACCCCTGTCTACTACAGGGCCCTCACCTCCACCCCTGTCACTACAGCTCACCTCTTA	2162
Db	1912	CTCCACCCCTGTCTACTACAGGGCCCTCACCTCCACCCCTGTCACTACAGCTCACCTCTTA	1971
QY	2163	CAGCTTTAATGCCAGGCCATGTCTGTCTCAAGGGCTCAAGACTTTCTAACTGGGA	2222
Db	1972	CAGCTTTAATGCCAGGGCCCATGTCTGTCTCAAGGGCTCAAGACTTTCTAACTGGGA	2031
QY	2223	TGTGTGTAGAGGAATCAAGGTACTTTGGGGGGCAAGAGACCTGTGTTATCTTCACAT	2282
Db	2032	TGTGTGTAGAGGAATGTAGAGTACCTTTGGGGGGCAAGAGACCTGTGTTATCTTCACAT	2091
QY	2283	CTACCCCTGTGACACTCACTGTGTGGACAGGATGTAAACAGACTTCCCGTGCATAAAAGG	2342
Db	2092	CTACCCCTGTGACACTCACTGTGTGGACAGGATGTAAACAGACTTCCCGTGCATAAAAGG	2151
QY	2343	TCACGCTCCCAACCCCGCCCTCCCTGCACCTCTGTCTCTCTCCAGTTTCATTCCTGG	2402
Db	2152	TCACGCTCCCAACCCCGCCCTCCCTGCACACTCTGTCTCTCTCCAGTTTCATTCCTGG	2211
QY	2403	AACACAGCAGGCGCAGGCAACCAATGGCCCCCAAAAGGCAGGAGATCCTCAGGCCCCAGC	2462
Db	2212	AACACAGCAGGCGCAGGCAACCAATGGCCCCCAAAAGGCAGGAGATCCTCAGGCCCCAGC	2271
QY	2463	CGCGGAGGCTGGAAGGGCTGGGAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGGTC	2522
Db	2272	CGCGGAGGCTGGAAGGGCTGGGAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGGTC	2331
QY	2523	TTTGTCTCTGTCCCGACAGCTCTGTGTAGACCAACGCGCAGATCACAGGGCACAGGCGAGA	2582
Db	2332	TTTGTCTCTGTCCCGACAGACTCTGTGTAGACCAACGCGCAGATCACAGGGCACAGGCGAGA	2391
QY	2583	GATAGTCTCTTTTGTGTCTTCTGGCCCTGTGGCTAGTGTAGTTTTCATAGCTTACAGT	2642
Db	2392	GATAGTCTCTTTTGTGTCTTCTGGCCCTGTGGCTAGTGTAGTTTTCATAGCTTACAGT	2451
QY	2643	ATTCGGCTTTGTACTGAGAAATAAAACACATTTTCTATAAAAAA	2702
Db	2452	ATTCGGCTTTGTACTGAGAAATAAAACACATTTTCTATAAAAAA	2511
QY	2703	AA 2704	
Db	2512	AA 2513	
RESULT 8			
ABLA0805			
ID ABLA0805 standard; cDNA: 2322 BP.			

AC	ABLA0805,	
XX		
DT	03-JUL-2002	(first entry)
XX		
DE	Human MAP kinase phosphatase-like enzyme encoding cDNA.	
XX		
KW	Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;	
KW	antihistaminic; antidiabetic; anorectic; cytostatic; cardiant; human;	
KW	antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;	
KW	neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;	
KW	antiallergic; dermatological; vulnerary; gene therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	
FT	location/Qualifiers	
FT	1..2322	
FT	/tag= a	
FT	/product= "MAP kinase phosphatase-like enzyme"	
FT	/note= "contains internal codon deletions"	
FT	/transl_except= "(pos: 156..157, aa: Leu)"	
FT	/note= "there is an apparent one codon deletion which	
FT	alters the reading frame"	
FT	/transl_except= "(pos: 180..181, aa: Arg)"	
FT	/note= "there is an apparent one codon deletion which	
FT	alters the reading frame"	
FT	/transl_except= "(pos: 414..415, aa: Val)"	
FT	/note= "there is an apparent one codon deletion which	
FT	alters the reading frame"	
FT	/transl_except= "(pos: 1197..1198, aa: Tyr)"	
FT	/note= "there is an apparent one codon deletion which	
FT	alters the reading frame"	
FT	/transl_except= "(pos: 2088..2089, aa: Xaa)"	
FT	/note= "there is an apparent one codon deletion which	
FT	alters the reading frame"	
XX		
PN	MO200220732-A2.	
XX		
PD	14-MAR-2002.	
XX		
XX	27-AUG-2001; 2001MO-EP09848.	
XX		
PR	07-SEP-2000; 2000US-230709P.	
XX		
PA	(FARB) BAYER AG.	
XX		
PI	Liou J;	
XX		
DR	WPI: 2002-339802/37.	
P	P-PSDB: ABB07845.	
XX		
PT	New human mitogen activated protein kinase phosphatase-like enzyme	
PT	polypeptide, regulators of which are useful for preventing, treating	
PT	allergies including asthma, diabetes, obesity, cancer and	
PT	cardiovascular diseases	
XX		
PS	Claim 1, Fig 13; 134pp. English.	
XX		
CC	The invention relates to a purified human mitogen activated protein (MAP)	
CC	kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed	
CC	by standard recombinant methodology. The MAP kinase phosphatase-like	
CC	enzyme and encoding polynucleotides are useful for screening for	
CC	modulators which are used for treating a MAP kinase phosphatase-like	
CC	enzyme dysfunction related disease, such as asthma, a central nervous	
CC	system disorder, diabetes, obesity, chronic obstructive pulmonary	
CC	disease, cancer or a cardiovascular disease. The enzyme can be regulated	
CC	to treat allergies including asthma, allergic rhinitis, atopic	
CC	dermatitis, and anaphylaxis, central nervous system disorders such as	
CC	brain injuries, Parkinson's disease, dementia, multiple sclerosis,	
CC	stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's	
CC	disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human	
CC	immunodeficiency virus (HIV) dementia, and cardiovascular diseases	
CC	including myocardial infarction, ischaemic diseases of the heart, atrial	
CC	and ventricular arrhythmia, hypertensive vascular diseases and peripheral	
CC		

	Query Match	Best Local Similarity	Score	DB	Length	2322
	Matches 1987; Conservative	90.9%;	Pred. No. 0;	Mismatches 75;	Indels 123;	Gaps 12.
CC	vascular diseases. The enzyme is useful in diagnostic assays for					
CC	detecting diseases and abnormalities or susceptibility to diseases or					
CC	abnormalities related to the presence of mutations in the encoding					
CC	nucleic acid sequences. The present sequence represents the human MAP					
XX	kinase phosphatase-like enzyme polypeptide encoding cDNA.					
XX	Sequence 2322 BP; 467 A; 751 C; 718 G; 386 T; 0 other:					
QY	132 CGGCGCTCCACGCGCGTGGGGCCCTGGGGACCAAGGCGTCCAGCGAAGAGTGCAGCTCCA	191				
DB	147 CCCCCACCCCTGGGGCTCTCTCTCGAGAGACCTGTCTCCAGCGAAGAGTGCAGCTCCA	206				
QY	192 GCGAAGCAGACACTTGGCGGTCTCC -GTGGGGCTGTCTCTGGAGTGCAGATGAGAGGG	250				
DB	207 GCGA----AGAGCTTGGCGGTCTCTCTGTGGGGCTGTCTCTGGAGTGCAGATGAGAGGG	262				
QY	251 ACNATGATGATGCAGCAGAGAGGCCAGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAAGG	310				
DB	263 ACAATGATGATGAGCAGCAGAGAGCCAGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAAGG	322				
QY	311 AGCTCCAGGGG---ACGAGACAGACTT---CGGGCAAGGATCCCAAGTCCCAAGAGA	365				
DB	323 AGCTCCAGGGGGAGCCACAGACAGACTTTCGGTGAGTCCCAAGTCCCAAGAGA	382				
QY	366 GAGAGCAGAGAGCAGCACCCTGCACCTCAGTACAGTCTGAGGCCAGAGATGACAT	425				
DB	383 GAGAGCAGAGAGCAGCACCCTGCACCTCAGTACAGTCTGAGGCCAGAGATGACAT	442				
QY	426 CCGCCCTGGCA--GCCACCTGAGGACCCCGGCTCTCCCGCTCCGCTACTGC-----	478				
DB	443 CCGACTTGGAGGCCCAAGCTGAGGACACCCCGGCTCTCCCGGATCCGGAATACCTTGCT	502				
QY	479 TGTGATTTCTAACAGAGAAG--GAGAAGGTCTGAG--CCAGATGAGACGGTCTCTGGG	536				
DB	503 TGTGATTTCTAACAGAGAAGGAGGAGGAGGTCTGAGGCCAGATGAGACGGTCTCTGGG	552				
QY	537 -CGTGATTTCCCTGACAGAGCTCCGCCAGCTGCAACCCGTGGGCTTGTTGCCCTCT	595				
DB	563 ACGTGGATTTCCCTGACAGAGCTCCGCCAGCTGCAACCCGTGGGCTTGTTGCCCTCT	622				
QY	596 GGAATGACACCCAGGTGTACTTAAATGAGAGACGGGGCTTCAGCTGACCTCTGTGTGGC	655				
DB	623 GGAATGACACCCAGGTGTACTTAAATGAGAGACGGGGCTTCAGCTGACCTCTGTGTGGC	682				
QY	656 AAAGCCGATCTTCAGGCCATCTCCATCCAGACATGTGGGACACATCCAGGATATTC	715				
DB	683 AAAGCCGATCTTCAGGCCATCTCCATCCAGACATGTGGGACACATCCAGGATATTC	742				
QY	716 ACCAAGATGTGAGGCGAGCTCTAGCCAGCGGCCCTTGAACGGGGTGGAGTCCCTCACT	775				
DB	743 ACCAAGATGTGAGGCGAGCTCTAGGACGCGGCCCTTGAACGGGGTGGAGTCCCTCACT	802				
QY	776 GGGCAGCCACTACAGAGAGAGACTGAATCCGAACAGAGCTGCCTCAATGATGACGG	835				
DB	803 GGGCAGCCACTACAGAGAGAGACTGAATCCGAACAGAGCTGCCTCAATGATGACGG	862				
QY	836 CTATGGCCAGCTTGAAGTCTCTGCGGCTCCCGAGCGGAGCTGTGGGGGTCTCCATAAC	895				
DB	863 CTATGGCCAGCTTGAAGTCTCTGCGGCTCCCGAGCGGAGCTGTGGGGGTCTCCATAAC	922				
QY	896 AGAGCAGATGAGACAGGCCATCCCTGCTACGCTGTGGAAGTGTTCAGTTCAGTAC	955				
DB	923 AGAGCAGATGAGACAGGCCATCCCTGCTACGCTGTGGAAGTGTTCAGTTCAGTAC	982				
QY	956 TGGAGAGTGTACTTCCAAAGAGATCCGCCAGGCTTGAAGCTGCGCTGGGGCTTCCCC	1015				
DB	983 TGGAGAGTGTGTCTTCCAGAGAGATCCGCCAGGCTTGAAGCTGCGCTGGGGCTTCCCC	1042				
QY	1016 TCCAGCAGTACCGTGAATCTCATGCAACAACGATGCTGCTGTGGACACAGCGGAGAC	1075				


```
|||||
Db 1043 TCACACATACCTGATCTTCATCGACAACAGATGCTGTGTGGACAGCGGGACC 1102
Qy 1076 GAGCCTCCGATCTTCCCAACCTCTACCTGGGCTCAGAGGGAAGCAGCAACCTG 1135
Db 1103 GAGCCTCCGATCTTCCCAACCTCTACCTGGGCTCAGAGGGAAGCAGCAACCTG 1162
Qy 1136 AGAGCTCAGAGGAACAGGGTCAACCACTTTGAACTGGCCGGGAGATTGCACT 1195
Db 1163 AGGAGCTCAGAGGAACAGGGTCAACCACTTTGAACTGGCCGGGAGATTGCACT 1219
Qy 1196 TCTACCTGAGCGCTTCACTACCAACAATGTGGCGCTTGGAGTAGAGAGTGGCCAGC 1255
Db 1220 TCTACCTGAGCGCTTCACTACCAACAATGTGGCGCTTGGAGTAGAGAGTGGCCAGC 1279
Qy 1256 TGTGCGCGCATCGGAAGAGAGCGACCGCTTCATGTAGAGCTGCAAGACACAGGCGCACCC 1315
Db 1280 TGTGCGCGCATCGGAAGAGAGCGACCGCTTCATGTAGAGCTGCAAGAGCACAGGCGCACCC 1339
Qy 1316 ACGTCTGTGTCACACTGCAAGATGGGCGTCAAGCGCGCTCAGCGGCCACAGTGTGCGCTATG 1375
Db 1340 ACGTCTGTGTCACACTGCAAGATGGGCGTCAAGCGCGCTCAGCGGCCACAGTGTGCGCTATG 1399
Qy 1376 CCATGAGACAGTACGATGACAGCTGAGAGAGCGCCCTGCGCCAGCTGCAAGAGCTCCGGC 1435
Db 1400 CCATGAGACAGTACGATGACAGCTGAGAGAGCGCCCTGCGCCAGCTGCAAGAGCTCCGGC 1459
Qy 1436 CCATGCGCGCGCGCCCAACCGCTGCTCGCGCGACGTCAGATACAGAGGCACTCCGCA 1495
Db 1460 CCATGCGCGCGCGCCCAACCGCTGCTCGCGCGACGTCAGATACAGAGGCACTCCGCA 1519
Qy 1496 CG----- 1497
Db 1520 CGGCGAGCGCGGACAGGACATGTCTGGAGCAAGAAGTGGGTCTCCCGACAGAGAGC 1579
Qy 1498 -----GCCAGACCTGAGGCTGTG 1517
Db 1580 ACCGAGCCCTGAAAGTCTCTACACCATTCGCCACCTCTCCGCGCAAGACCTGAGGCTGTG 1639
Qy 1518 GGGAGGAGAAAGTTGTAGACATGGAAGAGAGCGAGCGCCCGCAAGAAAGAGCTTGAGG 1577
Db 1640 GGGAGGAGAAAGTTGTAGACATGGAAGAGAGCGAGCGCCCGCAAGAAAGAGCTTGAGG 1699
Qy 1578 CACGGCCACGATATAACCTCCGAGGGGTATAGAGGTCCATCAGTCTTCCAGAGCCCTCT 1637
Db 1700 CACGGCCACGATATAACCTCCGAGGGGTATAGAGGTCCATCAGTCTTCCAGAGCCCTCT 1759
Qy 1638 TGGAGCTGGAGAGACCTCAGAGACAGTACATGCCAGAGGCTTCTCTCCACGAGT 1697
Db 1760 TGGAGCTGGAGAGACCTCAGAGACAGTACATGCCAGAGGCTTCTCTCCACGAGT 1819
Qy 1698 CTTACATGAGAGAGCTCTGACAGCCCTTCCACAGCTTGCAGAGCAAGAGGAGCGCAGC 1757
Db 1820 CTTACATGAGAGAGCTCTGACAGCCCTTCCACAGCTTGCAGAGCAAGAGGAGCGCAGC 1879
Qy 1758 AGGTGAGAGAGGGGCTCAGGCTGCGCTGAAGTCCCGGCACTCAGTGGTTACCTCCAGG 1817
Db 1880 AGGTGAGAGAGGGGCTCAGGCTGCGCTGAAGTCCCGGCACTCAGTGGTTACCTCCAGG 1939
Qy 1818 GCAGTGCCTGTGGCCAAACCGGAGCCCTTCCAGAGACAGAGCAAGAGGAGGGGCG 1877
Db 1940 GCAGTGCCTGTGGCCAAACCGGAGCCCTTCCAGAGACAGAGCAAGAGGAGGGGCG 1999
Qy 1878 AGGGGAGAGAGAGCGCTGCAATTTCTTACGCGCCAGGTTCCGGAAGGTGGTAGACAGG 1937
Db 2000 AGGGGAGAGAGAGCGCTGCAATTTCTTACGCGCCAGGTTCCGGAAGGTGGTAGACAGG 2059
Qy 1938 CCAGGCTCATGACAGTGGAGAGAGGCGGAGCCCTGAGCCCTCAGACATGCCACGCTC 1997
Db 2060 CCAGGCTCATGACAGTGGAGAGAGGCGC--GGCTGAGCCCTCAGACATGCCACGCTC 2117
Qy 1998 CCTGTACACTGAAGAGATCCCAACAACCTCTTGGAGAAACACCCCTACGCTGTGCGCGCA 2057
|||||
```

```
Db 2118 CCTGACACTGAAGAGATGCCAACACTCTTGGAGAAACACCCCTCAGCTGTGGCCGCA 2177
Qy 2058 CACATTCCTCTCAGCTCCGCCCATATCCGTCATCAGAGCCCTACCTCCACCCCTGCA 2117
Db 2178 CACATTCCTCTCAGCTCCGCCCATATCCGTCATCAGAGCCCTACCTCCACCCCTGCA 2237
Qy 2118 CTACGGCCTCAGCTCCACCCCTGTCACCTACAGAGCTTCACCTCTTACAGCCCTTAACTCCA 2177
Db 2238 CTACGGCCTCAGCTCCACCCCTGTCACCTACAGAGCTTCACCTCTTAACTCCA 2297
Qy 2178 GGCCCATGTCTGCTGTCCAAAGGC 2202
Db 2298 GGCCCATGTCTGCTGTCCAAAGGC 2322

RESULT 9
AAH14722
ID AAH14722 standard; cdna: 1755 BP.
XX
AC AAH14722;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cdna sequence SEQ ID NO:12452.
XX
KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PE 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
FA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12452; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
```

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other:

Query Match 49.4%; Score 1337; DB 22; Length 1755;
Best Local Similarity 94.7%; Pred. No. 1.9e-258;
Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

```

OY 1174 ATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTACCTACCAACAATGTGCGCTC 1233
DB 322 ATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTACCTACCAACAATGTGCGCTC 381
OY 1234 TGGGATGAGAGTGGGCGCGAGCTGCTGCCACATGGAAAGAGACGACCGCTCATTTGAG 1293
DB 382 TGGGATGAGAGTGGGCGCGAGCTGCTGCCACATGGAAAGAGACGACCGCTCATTTGAG 441
OY 1294 GCTGCAAGACACAGGACACCCACAGTGTGTGCTCCACTGCAAGATGGGCGCTCAGCGCTCA 1353
DB 442 GCTGCAAGACACAGGACACCCACAGTGTGTGCTCCACTGCAAGATGGGCGCTCAGCGCTCA 501
OY 1354 GCGGCCACAGTGTGCTGCTATGCTATGCAAGTACGATGCAAGTGTGCAAGGCGCTG 1413
DB 502 GCGGCCACAGTGTGCTGCTATGCTATGCAAGTACGATGCAAGTGTGCAAGGCGCTG 561
OY 1414 GCGCAGCTGAGAGAGTCCGCGCCATGCGCGCCCAACCTGGCTTCCGCGCGAGCTG 1473
DB 562 GCGCAGCTGAGAGAGTCCGCGCCATGCGCGCCCAACCTGGCTTCCGCGCGAGCTG 621
OY 1474 CAGATCTACACAGGAGCATCTGACGCGCCAGAACTGAGGGTGTGGGAGAGAGAAGTTGT 1533
DB 622 CAGATCTACACAGGAGCATCTGACGCGCCAGAACTGAGGGTGTGGGAGAGAGAAGTTGT 681
OY 1534 AGGATGGAAGAGAGCCAGGACGCGCGGAAAGAGAGCTTGGGCGACGCGCATATATAA 1593
DB 682 AGGATGGAAGAGAGCCAGGACGCGCGGAAAGAGAGCTTGGGCGACGCGCATATATAA 741
OY 1594 CCTCCGAGGGGTCATGAGTGCATGATCTTCTGAGAGCCCTTCTGGAGCTGGAGAGCAC 1653
DB 742 CCTCCGAGGGGTCATGAGTGCATGATCTTCTGAGAGCCCTTCTGGAGCTGGAGAGCAC 801
OY 1654 CTCAGAGACAGTACATGTCAGAGAGTCTTCTCTCCACAGAGTCTTCACATGAAGACC 1713
DB 802 CTCAGAGACAGTACATGTCAGAGAGTCTTCTCTCCACAGAGTCTTCACATGAAGACC 861
OY 1714 TCTGCAAGCTTCCACAGCTTGCAGAGACCAAGGAGGCGCAGCTGAGAGGGGCGC 1773
DB 862 TCTGCAAGCTTCCACAGCTTGCAGAGACCAAGGAGGCGCAGCTGAGAGGGGCGC 921
OY 1774 TCAGAGCTGCGCTGAAGTCCCGCCAGTACAGTGTATACCTCCAGGGCAGTCCCGTGGGCG 1833
DB 922 TCAGAGCTGCGCTGAAGTCCCGCCAGTACAGTGTATACCTCCAGGGCAGTCCCGTGGGCG 981
OY 1834 CAACCGGACCCAGGAGCTTCCAGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGCC 1893
DB 982 CAACCGGACCCAGGAGCTTCCAGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGCC 1041
OY 1894 CTGCAATTTCTCTACGCGCCAGGTTCCGGAAGGTGTGTGAGACAGGCGCAGCTGCATGACAG 1953
DB 1042 CTGCAATTTCTCTACGCGCCAGGTTCCGGAAGGTGTGTGAGACAGGCGCAGCTGCATGACAG 1101
OY 1954 TGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2013
DB 1102 TGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
OY 2014 GATCCACAGTCTTGGAGAGAACACCTCACTGTCTGTGTGCGGACACATCTCTCACT 2073
DB 1162 GATCCACAGTCTTGGAGAGAACACCTCACTGTCTGTGTGCGGACACATCTCTCACT 1221
OY 2074 CCGGCCATACCGGTACTACAGGCTCACTCCACCGCTGTACATAGGCGCTCACCTCC 2133
DB 1221 CCGGCCATACCGGTACTACAGGCTCACTCCACCGCTGTACATAGGCGCTCACCTCC 2181
```

```

DB 1222 CCGGCCATACCGGTACTACAGGCTCACTCCACCGCTGTACATAGGCGCTCACCTCC 1281
OY 2134 CACCCCTGTCTACTACAGGCTCACTCCACCGCTTAAGTCCAGGCCATGTGCGCTG 2193
DB 1282 CACCCCTGTCTACTACAGGCTCACTCCACCGCTTAAGTCCAGGCCATGTGCGCTG 1341
OY 2194 TCCAGGGCTCAAGACTTCTTAACCTGGGATGTGTGTAAGAGGAGTAAAGTACCTTTGGGG 2253
DB 1342 TCCAGGGCTCAAGACTTCTTAACCTGGGATGTGTGTAAGAGGAGTAAAGTACCTTTGGGG 1401
OY 2254 GCACAGCAGCCTAGTTTATTTCTACCTAGGCCCTGACACTCAGCTGTGGACGGAA 2313
DB 1402 GCACAGCAGCCTAGTTTATTTCTACCTAG----- 1434
OY 2314 TGAAGACAGAGTCCCGTGCAAAAAAGGGTCAAGGCTCCACCGCCCGCCCTCCCTGCA 2373
DB 1435 -----CCCTGGA 1441
OY 2374 CCTCTGTCTCTCCAGTTTCATCTCTGAAACAGCAGGCGCAGGCAACCAAGTGGCCCC 2433
DB 1442 CCTCTGTCTCTCCAGTTTCATCTCTGAAACAGCAGGCGCAGGCAACCAAGTGGCCCC 1501
OY 2434 AAGGAGGAGGAGATCTCAGGCGCCAGCGGGGGAGGCTGGAAGGCGTGGAGATCGCT 2493
DB 1502 AAGGAGGAGGAGATCTCAGGCGCCAGCGGGGGAGGCTGGAAGGCGTGGAGATCGCT 1561
OY 2494 TCCCTCATCCACTCCACCGGTCCAGGCTTGTGCTGCTCCCGACAGCTCTGTGACAC 2553
DB 1562 TCCCTCATCCACTCCACCGGTCCAGGCTTGTGCTGCTCCCGACAGCTCTGTGACAC 1621
OY 2554 CACGCGAGATCACAGGCGCACAGGCGCAGAGATAGTCTTTTGTCTTGTGCGCTCT 2613
DB 1622 CACGCGAGATCACAGGCGCACAGGCGCAGAGATAGTCTTTTGTCTTGTGCGCTCT 1681
OY 2614 GGCTAGTCACTTTTCTATAGCCCTTACAGTATCGCTTGTGCTGAGATATAACACAT 2673
DB 1682 GGCTAGTCACTTTTCTATAGCCCTTACAGTATCGCTTGTGCTGAGATATAACACAT 1741
OY 2674 TTTTCATA 2680
DB 1742 TTTTCATA 1748

RESULT 10
ID ABL40801
XX ABL40801 standard; DNA: 1755 BP.
XX
AC ABL40801:
XX
XX 03-JUL-2002 (first entry)
XX
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
XX Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
XX antidiabetic; antidiabetic; anorectic; cytosolic; cardiant; human;
XX antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
XX neuroleptic; anticonvulsant; anti-HIV; antirhythmic; hypotensive;
XX antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
XX WO200220732-A2.
XX
XX
PD 14-MAR-2002.
XX
XX
XX 27-AUG-2001; 2001WO-EP09848.
XX
XX 07-SEP-2000; 2000US-230709P.
XX
XX (FARB ) BAYER AG.
XX
XX Llou J;
XX
```


DR WPI: 2002-339802/37.
XX New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases -
PS Disclosure: Fig 4; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
XX
SQ Sequence 1755 BP: 350 A; 577 C; 505 G; 323 T; 0 other;
Query Match 49.4%; Score 1337; DB 24; Length 1755;
Best Local Similarity 94.7%; Pred. No. 1.9e-258;
Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;
QY 1174 ATGGCCCGGAGATTGACAACTTACCTGAGCGCTTACCTACACATGTGCGCTC 1233
Db 322 ATGGCCCGGAGATTGACAACTTACCTGAGCGCTTACCTACACATGTGCGCTC 381
QY 1234 TGGGATGAGAGTGGGCCACCTGCTGCCGACCTGGAAGAGACGACCGCTTATTAG 1293
Db 382 TGGGATGAGAGTGGGCCACCTGCTGCCGACCTGGAAGAGAGGACCGCTTATTAG 441
QY 1294 GCTCAAGAGACAGAGGACACCCAGCTGCTGTCACCTCAAGATGGGCGTACGCGCTCA 1353
Db 442 GCTCAAGAGACAGAGGACACCCAGCTGCTGTCACCTCAAGATGGGCGTACGCGCTCA 501
QY 1354 GCGGCCACAGTGTGGGCTATGCCATGAAGCAGTACGATGACAGCTGGAGCGCCCTG 1413
Db 502 GCGGCCACAGTGTGGGCTATGCCATGAAGCAGTACGATGACAGCTGGAGCGCCCTG 561
QY 1414 CGCCACGTGACAGAGCTCGCGCCCATGCGCCGCCAACCCTGGCTTCTGCGCAGCTG 1473
Db 562 CGCCACGTGACAGAGCTCGCGCCCATGCGCCGCCAACCCTGGCTTCTGCGCAGCTG 621
QY 1474 CAGATCTACAGAGGATCCCTGAGCGCAGACCTGAGGGGTGGGGGAGAAAGTTGT 1533
Db 622 CAGATCTACAGAGGATCCCTGAGCGCAGACCTGAGGGGTGGGGGAGAAAGTTGT 681
QY 1534 AGGATGGAAGAGAGCCAGGACGCCGAAAGAGAGCTGGGCGACGCGCATATAA 1593
Db 682 AGGATGGAAGAGAGCCAGGACGCCGAAAGAGAGCTGGGCGACGCGCATATAA 741
QY 1594 CCTCGAGAGGCTCATGAGTCTCATGAGTCTTCTGAGAGCCCTCTTGAAGCTGAGAGCAC 1653
Db 742 CCTCGAGAGGCTCATGAGTCTCATGAGTCTTCTGAGAGCCCTCTTGAAGCTGAGAGCAC 801
QY 1654 CTCAGAGACAGTACATGACAGAGGCTCTCTTCCACAGAGCTTACATGAAGAGCC 1713
Db 802 CTCAGAGACAGTACATGACAGAGGCTCTCTTCCACAGAGCTTACATGAAGAGCC 861

QY 1714 TCTGAGCCCTTCCACAGCTTGGACAGGACCAGAGGCCAGCAGGTGACAGGGGCC 1773
Db 862 TCTGAGCCCTTCCACAGCTTGGACAGGACCAGAGGCCAGCAGGTGACAGGGGCC 921
QY 1774 TCAGCTCCCTGAAGTCCCGCCAGCTAGTGTTAACCTCCAGGGCAGTGGCTGTGGC 1833
Db 922 TCAGCTCCCTGAAGTCCCGCCAGCTAGTGTTAACCTCCAGGGCAGTGGCTGTGGC 981
QY 1834 CAACCGAGCCAGGCTTCCAGAGACAGCAGGAGGAGGAGGAGGAGGAGGAGCC 1893
Db 982 CAACCGAGCCAGGCTTCCAGAGACAGCAGGAGGAGGAGGAGGAGGAGGAGCC 1041
QY 1894 CTGATTTCTCTACAGCCAGCTCCGGAAGTGTGAGACAGGCGCAGCTGATGACAG 1953
Db 1042 CTGATTTCTCTACAGCCAGCTCCGGAAGTGTGAGACAGGCGCAGCTGATGACAG 1101
QY 1954 TGGAGAGAGGGCCAGGCTTACAGCTTACACATGCCAGCTCCCTGACACTGAAGAG 2013
Db 1102 TGGAGAGAGGGCCAGGCTTACAGCTTACACATGCCAGCTCCCTGACACTGAAGAG 1161
QY 2014 GATCCACACTCTTGGAGAAACACCCCTCAGCTCTGTTGGCGCACATTCCTCTCAGCT 2073
Db 1162 GATCCACACTCTTGGAGAAACACCCCTCAGCTCTGTTGGCGCACATTCCTCTCAGCT 1221
QY 2074 CCGCCCATACCGCTCACTACAGCTCCTCCACACCCCTGTCACCTAGCGCTCAGCTCC 2133
Db 1222 CCGCCCATACCGCTCACTACAGCTCCTCCACACCCCTGTCACCTAGCGCTCAGCTCC 1281
QY 2134 CACCCCTGTACTACAGCTTACCTCTCTACAGCTTAACTCCAGGCCATGTCTGCTG 2193
Db 1282 CACCCCTGTACTACAGCTTACCTCTCTACAGCTTAACTCCAGGCCATGTCTGCTG 1341
QY 2194 TCAGAGGCTCAAGACTTCTTAACTGGGATGTGGTAGAGGAGTGAAGGTACCTTTGGG 2253
Db 1342 TCAGAGGCTCAAGACTTCTTAACTGGGATGTGGTAGAGGAGTGAAGGTACCTTTGGG 1401
QY 2254 GCAACAGCACCTAGTTTCACTTCAACTAGCCCTGACACACTGTCAGCAGGAA 2313
Db 1402 GCAACAGCACCTAGTTTCACTTCAACTAG----- 1434
QY 2314 TGAANAAGAGCTTCCGCTGCAAAAAGGCTACAGCTTCCACCCCGCCCTCCCTGCA 2373
Db 1435 -----CCCTGCA 1441
QY 2374 CCTCTGCTCTCTCCAGTTTATCTGGAACAGGCGCAGGCAACAGTGGGCC 2433
Db 1442 CCTCTGCTCTCTCCAGTTTATCTGGAACAGGCGCAGGCAACAGTGGGCC 1501
QY 2434 AAAGGACAGAGATCTCAGGCCCCAGCCGCGGAGGCTGGAAGGCTGCGAGATGCT 2493
Db 1502 AAAGGACAGAGATCTCAGGCCCCAGCCGCGGAGGCTGGAAGGCTGCGAGATGCT 1561
QY 2494 TCCCTCATCCACCTCCACCGGCTCCAGGCTTGTGCTGCTGTCGCCAGACTCTGTGACAC 2553
Db 1562 TCCCTCATCCACCTCCACCGGCTCCAGGCTTGTGCTGCTGTCGCCAGACTCTGTGACAC 1621
QY 2554 CAGGCGAGATCAGAGGAGCAGGCGCAGAGATAGCTCTTTTGTCTTCTGAGCTCT 2613
Db 1622 CAGGCGAGATCAGAGGAGCAGGCGCAGAGATAGCTCTTTTGTCTTCTGAGCTCT 1681
QY 2614 GCGTAGTCAGTTTTCATAGGCTTACAGTATCTGGCTTGTGACTGAGAAATAAACAT 2673
Db 1682 GCGTAGTCAGTTTTCATAGGCTTACAGTATCTGGCTTGTGACTGAGAAATAAACAT 1741
QY 2674 TTTTCATA 2680
Db 1742 TTTTCATA 1748
RESULT 11
ABNS9832
ID ABNS9832 standard; cdna; 2061 BP.
XX

AC ABNS9832;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 243.
XX
XX Human; antineoplastic; vulnery; antineoplastic; immunomodulator;
KW antineoplastic; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
XX 21-MAR-2002.
PD 10-SEP-2001; 2001WO-US26015.
PF 11-SEP-2000; 2000US-0659671.
PR (HYSE-) HYSEQ INC.
XX
PA Tang Y, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR MPI: 2002-292408/33.
DR P-PSDB: ABB97419.
XX
PT An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 243; 509bp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
XX
XX Sequence 2061 BP; 415 A; 672 C; 605 G; 369 T; 0 other;
SQ
Query Match 47.4%; Score 1280.8; DB 24; Length 2061;
Best Local Similarity 73.8%; Pred. No. 3.6e-247;
Matches 1995; Conservative 0; Mismatches 12; Indels 697; Gaps 1;
1 CGCTCTCTCGGCTCTGCGGCTCCAGAGCTGTCGCGGGGTTAGAGAGGCGCGTGC 60
DB 54 CGTCTCTCTCGGCTCTGCGGCTCCAGAGCTGTCGCGGGGTTAGAGAGGCGCGTGC 113
QY 61 CGGTGCGAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGCTCAGTGAACCTGTCG 120
DB 114 CGGTGCGAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGCTCAGTGAACCTGTCG 173
QY 121 CCCCCGGGACGCGCGCTTCACGCGCGTGGGCGCTGGAGCAGGCGGTCCGCAAG 180
DB 174 CCCCCGGGACGCGCGCTTCACGCGCGTGGGCGCTGGAGCAGGCGGTCCGCAAG 233
QY 181 AGTGCATCCAGGAGGAGGAGGAGCTTGGGCTGCTCGTGGGCGCTGCTGAGTGCAG 240
DB 234 AGTGCATCCAGGAGGAGGAGGAGCTTGGGCTGCTCGTGGGCGCTGCTGAGTGCAG 293
QY 241 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 294 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 353
QY 301 AGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

DB 354 AGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 413
QY 361 AAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 414 AAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 473
QY 421 GACATCCGCGCTGGAGCGCCAGCTGGAGGAGCGCGCGCTCCCGGCTCCGCTACCTG 480
DB 474 GACATCCGCGCTGGAGCGCCAGCTGGAGGAGCGCGCGCTCCCGGCTCCGCTACCTG 533
QY 481 GTAGTTTCTACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 534 GTAGTTTCTACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593
QY 541 GATTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 594 GATTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 653
QY 601 GACACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 654 GACACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 713
QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTTGGGCAACATCCAGATTTGACACCA 720
DB 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTTGGGCAACATCCAGATTTGACACCA 773
QY 721 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 774 GCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 833
QY 781 AGCCACTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 834 AGCCACTACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 893
QY 841 GCGGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 894 GCGGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 953
QY 901 CGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 954 CGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 992
QY 961 AGTGTACTTCCAAAGAGATCCGAGGCTGTGAGCTGCGGCTGGGCTCCCTCCAG 1020
DB 993 ----- 992
QY 1021 CAGTACCGTGAATGATGAGACACAGATGCTGCTGTTGGGCAACGCGGAGCCAGCC 1080
DB 993 ----- 992
QY 1081 TCCCGATCTTCCCGGCTTACCTGAGGCTGAGAGTGAAGGAGCAACCTGAGAGAG 1140
DB 993 ----- 992
QY 1141 CTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 993 ----- 992
QY 1201 CCTGAGCGCTTACCTACCAATGATGCGGCTCTGAGATGAGAGTGGCGCCAGCTGCTG 1260
DB 993 ----- 992
QY 1261 CCGCACTGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 993 ----- 992
QY 1321 CTGTGTCACTGCAAGATGGGCTCAGCGGCTGAGCGGCAACATGCTGCTATGCTATG 1380
DB 993 ----- 992
QY 1381 AAGCACTACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

```
Db 993 ----- 992
QY 1441 GCCGCCCAACCTGGCTTCCTGGCCAGCTGCAGATCTACAGAGGCACTCTGACGGC 1500
Db 993 ----- 992
QY 1501 AGAACCTGAGGTGTGGGAGAGAGAGTTGTAGGCATGGAAGAGAGCCAGAGCCCC 1560
Db 993 ----- 992
QY 1561 GAAGAAGAGCCTGGGCCACGGCCACGTATTAACCTCCGAGGGGTCATGAGTCCATCAG 1620
Db 993 ----- 992
QY 1621 TCTTCTGGAGCCCTCTTGAGAGCTGAGAGACGCTCAGAGACAGTGCAGATCCAGAGT 1680
Db 993 ----- 992
QY 1681 CTTCCTCTCCAGAGATCTTCACATGAAAGAGCCTCTGCAGCCCTTCCACACAGCTTGAAG 1740
Db 1037 CTTCCTCTCCAGAGATCTTCACATGAAAGAGCCTCTGCAGCCCTTCCACACAGCTTGAAG 1096
QY 1741 GACCAAGGAGGAGCCAGCAGGTGTGACAGAGGGGCTTACGCTGCTGAAGTCCGCCAGTC 1800
Db 1097 GACCAAGGAGGAGCCAGCAGGTGTGACAGAGGGGCTTACGCTGCTGAAGTCCGCCAGTC 1156
QY 1801 AGTGTTCACCTCCAGAGGAGGTGGCTGGTGGCCAAACCGGACCCAGGCTTCCAGAGCA 1860
Db 1157 AGTGTTCACCTCCAGAGGAGGTGGCTGGTGGCCAAACCGGACCCAGGCTTCCAGAGCA 1216
QY 1861 GGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGACATTTCTCTACGCCAGGTTCCG 1920
Db 1217 GGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGACATTTCTCTACGCCAGGTTCCG 1276
QY 1921 GAAGGTGTGAGACAGAGCCAGCTGTACATGACAGTGGAGAGAGGCGAGGCTTGAGCCT 1980
Db 1277 GAAGGTGTGAGACAGAGCCAGCTGTACATGACAGTGGAGAGAGGCGAGGCTTGAGCCT 1336
QY 1981 CACACATGCCCCAGCTCCCTGACACTGAAGAGATCCCAACTCTTGAGAGAAACACC 2040
Db 1337 CACACATGCCCCAGCTCCCTGACACTGAAGAGATCCCAACTCTTGAGAGAAACACC 1396
QY 2041 TCACGTCTGTTGCGCAGACATTCCTCTCAGCTCCGCCCATACCCGTACTACAGCCTC 2100
Db 1397 TCACGTCTGTTGCGCAGACATTCCTCTCAGCTCCGCCCATACCCGTACTACAGCCTC 1456
QY 2101 ACCTCCACACCTGTCTACTACGGCTCACCTCCACCCCTGTCTACTACAGCCTCACCTC 2160
Db 1457 ACCTCCACACCTGTCTACTACGGCTCACCTCCACCCCTGTCTACTACAGCCTCACCTC 1516
QY 2161 TACAGCCTTAAGTCCAGAGCCCATGTGCTGCTCAAGAGGCTCAAGACTTCTTAAGTGG 2220
Db 1517 TACAGCCTTAAGTCCAGAGCCCATGTGCTGCTCAAGAGGCTCAAGACTTCTTAAGTGG 1576
QY 2221 GATGTGTAGAGGAGTGAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCATTTCTCAA 2280
Db 1577 GATGTGTAGAGGAGTGAAGGTACCTTTGGGGGCAACAGCACCCTAGTTTCATTTCTCAA 1636
QY 2281 CTCTAGCCCTGCACACTACCTGTGTGGCAGCGAATATAAAGAGAGCTTCCCGTCAAAAAG 2340
Db 1637 CTCTAGCCCTGCACACTACCTGTGTGGCAGCGAATATAAAGAGAGCTTCCCGTCAAAAAG 1696
QY 2341 GGTTCAGGCTCCACAGCCCGCCCTCTCTGCACTCCCTGCTCTCTCCAGTTCATTTCT 2400
Db 1697 GGTTCAGGCTCCACAGCCCGCCCTCTCTGCACTCCCTGCTCTCTCCAGTTCATTTCT 1756
QY 2401 GGAACACAGCAGGAGCAGCAACCACTGGGCCCCCAAGAGGAGGAGAGATCTTAGGCCCA 2460
Db 1757 GGAACACAGCAGGAGCAGCAACCACTGGGCCCCCAAGAGGAGGAGAGATCTTAGGCCCA 1816
QY 2461 GCGCGGGAGGCTGGAAGGCTGCGAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGG 2520
Db 1817 GCGCGGGAGGCTGGAAGGCTGCGAGATCGCTTCCCTCATCCACCTCCACCGGTCCAGG 1876
```

```
QY 2521 TCTTTGCTGTGCTCCCAAGACCTCTGTGACACACAGGCGAGATCAGAGGCAACAGGCCA 2580
Db 1877 TCTTTGCTGTGCTCCCAAGACCTCTGTGACACACAGGCGAGATCAGAGGCAACAGGCCA 1936
QY 2581 GAGATAGTCTTCTTTTGTGCTTTCTGGCCCTCTGCTAGTCAAGTTTTCATAGCTTACA 2640
Db 1937 GAGATAGTCTTCTTTTGTGCTTTCTGGCCCTCTGCTAGTCAAGTTTTCATAGCTTACA 1996
QY 2641 GTATCTGCTTTGTACTGAGAAATATAACACATTTTCATATAAAAAAAAAAAAAA 2700
Db 1997 GTATCTGCTTTGTACTGAGAAATATAACACATTTTCATATTGTATTTCATAAAAAA 2056
QY 2701 AAAA 2704
Db 2057 AAAA 2060

RESULT 12
AAF63578
ID AAF63578 standard; cDNA: 1026 BP.
AC AAF63578;
XX
XX
XX 11-MAY-2001 (first entry)
DE Human phosphatase NP_060746_h coding sequence.
XX
XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-LeFebvre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamatoma; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M.
PI WPI: 2001-211226/21.
XX
XX P-PSDB: AAB73226.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
XX Claim 1; Fig 4; 138pp: English.
XX
XX The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase coding sequence.
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-LeFebvre
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamatomas.
XX
XX Sequence 1026 BP: 217 A; 324 C; 326 G; 159 T; 0 other;
```

Query Match 30.3%; Score 820; DB 22; Length 1026;
 Best Local Similarity 90.4%; Pred. No. 6.5e-155;
 Matches 928; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY 1048 ATGCTGCTGCTGGTGGCAGACGGGAGCCCTCCGCACTTCTCCCGCACTCTACTG 1107
 |||||||
 Db 1 ATGCTGCTGCTGGTGGCAGACGGGAGCCCTCCGCACTTCTCCCGCACTCTACTG 60
 QY 1108 GGGTCAGAGTGGAAACGAGCAACCTGGAGAGCTGAGAGGAAAGGGTACCCACATC 1167
 |||||||
 Db 61 GGGTCAGAGTGGAAACGAGCAACCTGGAGAGCTGAGAGGAAAGGGTACCCACATC 120
 QY 1168 TTGAACATGCGCGGAGATTGACAACTTCTACCTGAGGCTTTCACCTACCAATGTG 1227
 |||||||
 Db 121 TTGAACATGCGCGGAGATTGACAACTTCTACCTGAGGCTTTCACCTACCAATGTG 180
 QY 1228 CGCCTCTGGGATGAGAGTGGCGCCACACTGCTGCCACTGGAAGAGACGACCGCTTC 1287
 |||||||
 Db 181 CGCCTCTGGGATGAGAGTGGCGCCACACTGCTGCCACTGGAAGAGACGACCGCTTC 240
 QY 1288 ATTGAGGCTGCAAGAGCAGAGGACCCAGCTGCTGCTCCACTGCAAGATGGGCTGAGC 1347
 |||||||
 Db 241 ATTGAGGCTGCAAGAGCAGAGGACCCAGCTGCTGCTCCACTGCAAGATGGGCTGAGC 300
 QY 1348 CGCTCAGCGGCGCAGAGTGTGCTGCTATGCCATGAGCAATGCAAGTACAGCTTGAGCAG 1407
 |||||||
 Db 301 CGCTCAGCGGCGCAGAGTGTGCTGCTATGCCATGAGCAATGCAAGTACAGCTTGAGCAG 360
 QY 1408 GGCCTGGCGCCAGCGTGGAGAGCTCGGGCCATGGCGCCGCAACCTGGCTTCTGCGC 1467
 |||||||
 Db 361 GGCCTGGCGCCAGCGTGGAGAGCTCGGGCCATGGCGCCGCAACCTGGCTTCTGCGC 420
 QY 1468 CAGCTGAGATCTACAGAGGCACTCTGACG----- 1497
 |||||||
 Db 421 CAGCTGAGATCTACAGAGGCACTCTGACGCGCCAGCCGCAAGCCATGTCTGGAGCAG 480
 QY 1498 ----- 1497
 |||||||
 Db 481 AAAGTGGTGGGTCTCCCGCAGAGAGACCCAGCCCTGAAGTCTCTACACATTCCCA 540
 QY 1498 -----GCCAAGACCTGAGGGTGTGGGAGAGAAAGTTGAGCATGGAAGAGAGC 1549
 |||||||
 Db 541 CCTCTTCCGCGAGACCTGAGGGTGTGGGAGAGAAAGTTGAGCATGGAAGAGAGC 600
 QY 1550 CAGCAGCGCCCGAAAGAGAGCTGGGCGCAGCGCACGTATAAAGCTCCGAGGGGTATG 1609
 |||||||
 Db 601 CAGCAGCGCCCGAAAGAGAGCTGGGCGCAGCGCACGTATAAAGCTCCGAGGGGTATG 660
 QY 1610 AGGTCCATCACTTCTGAGAGCCCTCTTGAGAGCTGAGAGCACCTCAGAGACCACTGAC 1669
 |||||||
 Db 661 AGGTCCATCACTTCTGAGAGCCCTCTTGAGAGCTGAGAGCACCTCAGAGACCACTGAC 720
 QY 1670 ATGTCAGAGGCTTCTCTCCAGAGAGCTTCAATGAAGAGCCCTGAGAGCCCTTCCA 1729
 |||||||
 Db 721 ATGTCAGAGGCTTCTCTCCAGAGAGCTTCAATGAAGAGCCCTGAGAGCCCTTCCA 780
 QY 1730 CAGCTTGAAGAGCAAGAGAGGAGAGAGTGAAGAGGGGGCTCAGGCTCGCCCTGAAG 1789
 |||||||
 Db 781 CAGCTTGAAGAGCAAGAGAGGAGAGAGTGAAGAGGGGGCTCAGGCTCGCCCTGAAG 840
 QY 1790 TCCCGCCAGTCACTGTTACCTCCAGAGGAGTGCCTGTTGGCCAAACCGGACAGGCC 1849
 |||||||
 Db 841 TCCCGCCAGTCACTGTTACCTCCAGAGGAGTGCCTGTTGGCCAAACCGGACAGGCC 900
 QY 1850 TTCCAGAGGAGAGAGAGGGGAGAGGGGAGAGGGGAGAGCCCTGATTTCTCTAAG 1909
 |||||||
 Db 901 TTCCAGAGGAGAGAGAGGGGAGAGGGGAGAGGGGAGAGCCCTGATTTCTCTAAG 960
 QY 1910 CCCAGGTTCCGAAAGTGTGAGACAGGCGACGCTGATGACAGTGAAGAGAGGGGAG 1969
 |||||||
 Db 961 CCCAGGTTCCGAAAGTGTGAGACAGGCGACGCTGATGACAGTGAAGAGAGGGGAG 1020

QY 1970 GCCTGA 1975
 |||||||
 Db 1021 GCCTGA 1026

RESULT 13

ID ABL40803 standard; DNA; 599 BP.

XX ABL40803;

DT 03-JUL-2002 (first entry)

XX Human MAP kinase phosphatase-like enzyme DNA fragment.

XX Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
 KW antidiabetic; anorectic; cytosolic; cardiant; human;
 KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
 KW neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
 KW antiallergic; dermatological; vulnerary; gene therapy; ds.

XX Homo sapiens.

PN MO200220732-A2.

XX 14-MAR-2002.

PF 27-AUG-2001; 2001WO-EP09848.

PR 07-SEP-2000; 2000US-230709P.

PA (FARB) BAYER AG.

PI Liou J;

DR WPI; 2002-339802/37.

PT New human mitogen activated protein kinase phosphatase-like enzyme
 PT polypeptide, regulators of which are useful for preventing, treating
 PT allergies including asthma, diabetes, obesity, cancer and
 PT cardiovascular diseases

PS Disclosure; Fig 8; 134pp; English.

XX The invention relates to a purified human mitogen activated protein (MAP)
 CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
 CC by standard recombinant methodology. The MAP kinase phosphatase-like
 CC enzyme and encoding polynucleotides are useful for screening for
 CC modulators which are used for treating a MAP kinase phosphatase-like
 CC enzyme dysfunction related disease, such as asthma, a central nervous
 CC system disorder, diabetes, obesity, chronic obstructive pulmonary
 CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
 CC to treat allergies including asthma, allergic rhinitis, atopic
 CC dermatitis, and anapylaxis; central nervous system disorders such as
 CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
 CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
 CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
 CC including myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
 CC vascular diseases. The enzyme is useful in diagnostic assays for
 CC detecting diseases and abnormalities or susceptibility to diseases or
 CC abnormalities related to the presence of mutations in the encoding
 CC nucleic acid sequences. The present sequence represents the human MAP
 CC kinase phosphatase-like enzyme DNA fragment.

XX Sequence 599 BP; 135 A; 185 C; 183 G; 96 T; 0 other;

Query Match 22.2%; Score 599; DB 24; Length 599;
 Best Local Similarity 100.0%; Pred. No. 1.1e-110;
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1117 TGGACGCGACCAACTGGAGAGCTGACAGAGAGAGGTCACCCACATCTTGAACATG 1176

```
|||||
Db 1 TCGACGACGCAACCTGGAGAGCTGCAGAGAACAGGCTCCACCATCTTGAACATG 60
OY 1177 GCCCGGAGATTGACACTTACCTTGAGCGCTTACCTACCACAAATGGCGCTTGG 1236
Db 61 GCCCGGAGATTGACACTTACCTTGAGCGCTTACCTACCACAAATGGCGCTTGG 120
OY 1237 GATGAGAGTGGCGCCAGCTGCTGCCGACTGGAAGAGACGACCGCTTCAATTAGAGCT 1296
Db 121 GATGAGAGTGGCGCCAGCTGCTGCCGACTGGAAGAGACGACCGCTTCAATTAGAGCT 180
OY 1297 GCAAGAGCACAGGAGCAACCCAGCTGCTGCTCACTGCAGAGATGGCGTCAAGCCCTCAGCG 1356
Db 181 GCAAGAGCACAGGAGCAACCCAGCTGCTGCTCACTGCAGAGATGGCGTCAAGCCCTCAGCG 240
OY 1357 GCCAGAGTGGCTGCTATGCGCTAGTGAAGACAGTACGATGAGCTGGAGCAGGCGCTGGCG 1416
Db 241 GCCAGAGTGGCTGCTATGCGCTAGTGAAGACAGTACGATGAGCAGGCGCTGGCG 300
OY 1417 CACGTGACAGAGCTCCGCGCCATCCGCCCCCAACCCCTGCTTCTGCGCAGCTGCAG 1476
Db 301 CACGTGACAGAGCTCCGCGCCATCCGCCCCCAACCCCTGCTTCTGCGCAGCTGCAG 360
OY 1477 ATCTACACAGGCGATCCTGACAGCGCAAACTGAGGGTGGTGGAGAGAGAGTGTAGG 1536
Db 361 ATCTACACAGGCGATCCTGACAGCGCAAACTGAGGGTGGTGGAGAGAGAGTGTAGG 420
OY 1537 CATGGAAGAGACGCGGAGCAACCCGGAAGAAGAGCTGGCGCCAGCGCATATAACT 1596
Db 421 CATGGAAGAGACGCGGAGCAACCCGGAAGAAGAGCTGGCGCCAGCGCATATAACT 480
OY 1597 CCGAGGGGTCATGAGTGCATCATGCTTCTGAGGCCCTCTTGGAGCTGAGAGCACTC 1656
Db 481 CCGAGGGGTCATGAGTGCATCATGCTTCTGAGGCCCTCTTGGAGCTGAGAGCACTC 540
OY 1657 AGAGACCACTGACATGCCAGAGAGGTTCTTCTTCCACAGAGTCTTACATGAAGACCTC 1715
Db 541 AGAGACCACTGACATGCCAGAGAGGTTCTTCTTCCACAGAGTCTTACATGAAGACCTC 599

RESULT 14
ABL40802
ID ABL40802 standard; DNA; 409 BP.
XX
AC ABL40802;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antidiabetic; anorectic; cytosolic; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antidiabetic; hypotensive;
KW antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO2002/20732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB ) BAYER AG.
XX
PI Liou J.
XX
DR WPI: 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
```

```
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
PS Disclosure: Fig 7; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
SQ Sequence 409 BP; 87 A; 127 C; 121 G; 73 T; 1 other:
XX
Query Match 14.7%; Score 397; DB 24; Length 409;
Best Local Similarity 99.5%; Pred. No. 3e-70;
Matches 408; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 912 GCGATCCGCTGCTGACGCTGGAAGTGTGATGTCAGTGCCTGAGAGTGCATCTC 971
Db 1 GCGATCCGCTGCT-AGCTGTGGAAGTGTGATGTCAGTGCCTGAGAGTGCATCTC 59
OY 972 CAAAGATCCGCGCAGGCTGAGCTGGGCTGGGCGCCCTCCAGCACTACCGTGA 1031
Db 60 CAAAGATCCGCGCAGGCTGAGCTGGGCTGGGCGCCCTCCAGCACTACCGTGA 119
OY 1032 CTTTCATGCACACAGATGCTGCTGCTGCTGACAGCGGAGCCCTCCGCTCTT 1091
Db 120 CTTTCATGCACACAGATGCTGCTGCTGCTGACAGCGGAGCCGAGCTCCGCTATCTT 179
OY 1092 CCCCACCTTACCTGCGCTCAGAGTGAAGCAGCAAACTGAGAGCTGCAGAGAA 1151
Db 180 CCCCACCTTACCTGCGCTCAGAGTGAAGCAGCAAACTGAGAGCTGCAGAGAA 239
OY 1152 CAGGCTACCCCATCTGAACATGCGCGGAGATGACAACTTACCTCCGAGCGCTT 1211
Db 240 CAGGCTACCCCATCTGAACATGCGCGGAGATGACAACTTACCTCCGAGCGCTT 299
OY 1212 CACCTACCAATGTGCCCTTGGATGAGAGTGCAGCCAGCTGCTGCAGCTGAA 1271
Db 300 CACCTACCAATGTGCCCTTGGATGAGAGTGCAGCCAGCTGCTGCAGCTGAA 359
OY 1272 GGAGAGCAGCGCTTATGAGCGCTGCAAGAGCAGCGCACCCAGCTGC 1321
Db 360 GGAGAGCAGCGCTTATGAGCGCTGCAAGAGCAGCGCACCCAGCTGC 409

RESULT 15
AAH98183/C
ID AAH98183 standard; cDNA; 717 BP.
XX
AC AAH98183;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST-derived coding sequence SEQ ID NO: 40.
XX
```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:54:09 ; Search time 13643.8 Seconds
(without alignments) 5767.767 Million cell updates/sec

Title: US-09-761-640-1
Sequence: 1 cgcctccctgcgtcgtcg9.....aaaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2583.2	95.8	2808	9 AK074432	AK074432 Homo sapi
2	2562.8	94.5	2781	9 AK099939	AK099939 Sequence
3	2352.8	87.0	2905	9 AK094226	AK094226 Homo sapi
4	1774	65.6	1980	6 AX451343	AX451343 Sequence
5	1699	62.8	2322	6 AX398774	AX398774 Sequence
6	1665.2	61.6	1905	9 BC004176	BC004176 Homo sapi
7	1665.2	61.6	1905	9 BC004210	BC004210 Homo sapi
8	1416	52.4	1416	6 AX451362	AX451362 Sequence
9	1391.4	51.5	1416	6 AB072360	AB072360 Homo sapi
10	1337	49.4	1755	6 AX398768	AX398768 Sequence
11	1337	49.4	1755	9 AK001790	AK001790 Homo sapi
12	1285.6	47.5	2064	9 AK000522	AK000522 Homo sapi
13	1280.8	47.4	2061	6 AX405828	AX405828 Sequence
14	1155.2	42.7	2736	10 BC028922	BC028922 Mus muscu
15	1004	37.1	150903	2 AP002776	AP002776 Homo sapi
16	1004	37.1	171566	2 AP001885	AP001885 Homo sapi
17	820	30.3	1026	6 AX086034	AX086034 Sequence
18	599	22.2	599	6 AX398772	AX398772 Sequence
19	575.2	21.3	580	9 HUMY153C10	HUMY153C10 Homo sapi
20	397	14.7	409	6 AX398771	AX398771 Sequence
21	358.4	13.3	2280	6 AB072356	AB072356 Homo sapi
22	358.4	13.3	2433	9 AK095421	AK095421 Homo sapi
23	358.4	13.3	3817	9 AB072355	AB072355 Homo sapi
24	358.4	13.3	6374	6 AX180874	AX180874 Sequence
25	336.8	12.5	1949	6 AX223960	AX223960 Sequence
26	336	12.4	426	6 AX398765	AX398765 Sequence
27	313.4	11.6	1711	6 AX223964	AX223964 Sequence
28	313.2	11.6	334	6 AX337360	AX337360 Sequence
29	306.8	11.3	1796	9 AB072358	AB072358 Homo sapi
30	297.2	11.0	3488	6 AX406972	AX406972 Sequence
31	293.6	10.9	4417	3 AB036834	AB036834 Drosophill
32	287.4	10.6	2260	6 AX180876	AX180876 Sequence
33	279.4	10.3	1771	6 AX202239	AX202239 Sequence
34	241	8.9	494	6 AX398773	AX398773 Sequence
35	233.8	8.6	279	11 G22634	G22634 human STS W
36	208.6	7.7	1052	6 AX223966	AX223966 Sequence
37	142.8	5.3	386	9 AF484838	AF484838 Homo sapi
38	142	5.3	113474	2 AC013932	AC013932 Drosophill
39	142	5.3	181132	3 AC008206	AC008206 Drosophill
40	142	5.3	227398	3 AE003750	AE003750 Drosophill
41	121.4	4.5	5463	9 AB037719	AB037719 Homo sapi
42	104.4	3.9	1390	6 AX237020	AX237020 Sequence
43	103.4	3.8	531	6 AX237022	AX237022 Sequence
44	101	3.7	1031	9 AB072357	AB072357 Homo sapi
45	98.8	3.7	1236	6 AX354534	AX354534 Sequence

ALIGNMENTS

RESULT 1

AK074432

LOCUS AK074432 2808 bp mRNA linear PRI 15-FEB-2002

DEFINITION Homo sapiens cDNA FLJ23852 fls, clone KAT12021.

ACCESSION AK074432

VERSION AK074432.1 GI:1867703

KEYWORDS oligo capping, fls (full insert sequence).

SOURCE Homo sapiens signet-tring cell carcinoma cell_line:KATO III CDNA to mRNA, clone:lib:KAT clone:KAT12021.

ORGANISM Homo sapiens

REFERENCE 1 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,

TITLE Nakamura, Y., Isogai, T. and Sugano, S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE 2 (bases 1 to 2808)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shihahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@leims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan. cDNA full insert sequencing:
Research Association for Biotechnology: cDNA library construction,
5'- & 3'-end one pass sequencing: Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).

FEATURES

source

CDS

BASE COUNT	585 a	912 c	823 g	488 t
ORIGIN				

Query Match	95.5%;	Score 2583.2;	DB 9;	Length 2808;
Best Local Similarity	96.28;	Pred. No. 0;		
Matches 2696; Conservative	0;	Mismatches 8;	Indels 98;	Gaps 1.

[illegible]

Db	304	AGTAGGAGGAGGCTCCACGGGGACACGACAGACTTGGGCAAGGATCCGAGTCCCGAG	363
Oy	361	AAGCAGAGAGACGACGAGGACGACCTTCGACCTCATGATGTACACTGCTGAGGCCGAGAT	420
Db	364	AAGCAGAGAGAGAGAGAGGCGAGCACTCGACCTCATGATGTACACTGCTGAGGGCCGAGAT	423
Oy	421	GACATCCGGCTGGGAGCGCCAGCTGAGAGGACCCCGGCTCCCGGCTCGGTAAGTGTG	480
Db	424	GACATCCGGCTGGGAGCGCCAGCTGAGAGGACCCCGGCTCCCGGCTCGGTAAGTGTG	483
Oy	481	GTAGTTCCTACACGAGAGAGAGAGGTTGAGCGAGAGATGAGACGGTCTCTGGGCGTG	540
Db	484	GTAGTTCCTACACGAGAGAGAGAGGTTGAGCGAGAGATGAGACGGTCTCTGGGCGATG	543
Oy	541	GATTTCCTGACAGCACTCTCCCGACGTGACCTGGGCTGGTCTGTCGGCCCTGGAGT	600
Db	544	GATTTCCTGACAGCACTCTCCCGACGTGACCTGGGCTGGTCTGTCGGCCCTGGAGT	603
Oy	601	GACACCCAGGTGTACTTAGTGGAGAGACGGGGCTTCAGCGTAGCTGTGGTGGCAAGC	660
Db	604	GACACCCAGGTGTACTTAGTGGAGAGAGCGGGGCTTCAGCGTAGCTGTGGTGGCAAGC	663
Oy	661	CGGATCTTCAAGCCCATCTCCATTCAGACCATGTGGGCCACACTCCAGGTATATGACCAA	720
Db	664	CGGATCTTCAAGCCCATCTCCATTCAGACCATGTGGGCCACACTCCAGGTATATGACCAA	723
Oy	721	GCAATGTAGAGCAGCTCTAGGCAAGCGGCTTGTACCGGGTGGGAGTGGCCCTCACTGGGCC	780
Db	724	GCAATGTAGAGAGCTCTAGGCAAGCGGCTTGTACCGGGTGGGAGTGGCCCTCACTGGGCC	783
Oy	781	AGCCACTTACAGAGAGAGACTCGGAACCTCCGAACAGAGCTGCTCAATGATGTGAGCGGCTATG	840
Db	784	AGCCACTTACAGAGAGAGACTCGGAACCTCCGAACAGAGCTGCTCAATGATGTGAGCGGCTATG	843
Oy	841	GCCACCTGGAGTCTTGCGGGCTCCGAGCGCCGAGCTGGGGGCTCCACAGAACAGAG	900
Db	844	GCCACCTGGAGTCTTGCGGGCTCCGAGCGCCGAGCTGGGGGCTCCACAGAACAGAG	903
Oy	901	CAGATGAGACGAGCGGATCCGTGCTGATGTGGAAAGTGTGATGTACATGACCTGGAG	960
Db	904	CAGATGAGACGAGCGGATCCGTGCTGATGTGGAAAGTGTGATGTACATGACCTGGAG	963
Oy	961	AGTGCACTTCCAAAGAGATCCGCGAGGCTGTGAGACTCGGCTGGGGCTCCCGCTCCAG	1020
Db	964	AGTGCACTTCCAAAGAGATCCGCGAGGCTGTGAGACTCGGCTGGGGCTCCCGCTCCAG	1023
Oy	1021	CAGTACCGTGACTTCATCGAACCAACAGAGATGTCGCTGTGTGGCACAGCGGGACCGACC	1080
Db	1024	CAGTACCGTGACTTCATCGAACCAACAGAGATGTCGCTGTGTGGCACAGCGGGACCGACC	1083
Oy	1081	TCCCGCATCTTCCCGCAACCTACACGTGGGCTCAGAGTGGAGAACGCAACACTGGAGAG	1140
Db	1084	TCCCGCATCTTCCCGCAACCTACACGTGGGCTCAGAGTGGAGAACGCAACACTGGAGAG	1143
Oy	1141	CTGCGAGAGAACAGGGTACCCACATTTAAMATGGCCCGGAGATTTACAACTTCTAC	1200
Db	1144	CTGCGAGAGAACAGGGTACCCACATTTAAMATGGCCCGGAGATTTACAACTTCTAC	1203
Oy	1201	CCTGAGCGCTTCACCTTACCAATGTGCGCTCTTGAGATBAGAGTCTGGGCCACGCTGCTG	1260
Db	1204	CCTGAGCGCTTCACCTTACCAATGTGCGCTCTTGAGATBAGAGTCTGGGCCACGCTGCTG	1263
Oy	1261	CCGACCTGGAGGAGACGACCGCTTCATTTGAGAGGTGCAAGAGACACAGGGCACCCACGG	1320
Db	1264	CCGACCTGGAGGAGAGACGACCGCTTCATTTGAGAGGTGCAAGAGACACAGGGCACCCACGG	1323
Oy	1321	CTGGTCCACTCAGATGTGGGCTGAGCCGCTCAGCGGACCAAGTGTGTGGCTATATGCCATG	1380
Db	1324	CTGGTCCACTCAGATGTGGGCTGAGCCGCTCAGCGGACCAAGTGTGTGGCTATATGCCATG	1383
Oy	1381	AAGCAGTACGAATCAGCTTGAGAGAGGCTTCGCCACGTGTGAGAGAGTCCGGCCCATC	1440
Db	1384	AAGCAGTACGAATCAGCTTGAGAGAGGCTTCGCCACGTGTGAGAGAGTCCGGCCCATC	1443

Oy	1441	GGCCGCCCAACCTGGCTTCCTCGCCCGACCTGACAGATTTACAGGGCATCTTGACG---	1497
Db	1444	GGCCGCCCAACCTGGCTTCCTCGCCCGACGTCGACATTTACAGGGCATCTTGACG	1503
Oy	1498	-----	1497
Db	1504	AGCCCGCAGAGCCATGCTGTGGAGCAGAAAGTGGGTGGGTCTCCACAGAGAGACCCA	1563
Oy	1498	-----GCCAAGCCGAGAGGTGGGGAG	1562
Db	1564	GGCCCTGAAGTCTCTACACCATTCCCACTCTTCCGCCAAGAACTGAGAGGTGGGGAG	1633
Oy	1523	GAGAAAGTTGTAGGCATGGAAGAGACCCAGCAGCCCGCAAAAGAAAGAGCCTGGGCCACG	1582
Db	1624	GAGAAAGTTGTAGGCATGGAAGAGACCCAGCAGCCCGCAAAAGAAAGAGCCTGGGCCACG	1683
Oy	1583	CCACGTTAAACCTCCGAGGGGTCATGAGGTCCATCAGTCTTCTGAGACCTCTTGGAG	1642
Db	1684	CCACGTTAAACCTCCGAGGGGTCATGAGGTCCATCAGTCTTCTGAGACCTCTTGGAG	1743
Oy	1643	CTGAGAGACACTCAGAGACACAGTGCATGACGAGAGGTCTTCTCTCCACAGACTTCA	1702
Db	1744	CTGAGAGACACTCAGAGACACAGTGCATGACGAGAGGTCTTCTCTCCACAGACTTCA	1803
Oy	1703	CATGAGAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGAGGTG	1762
Db	1804	CATGAAAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGAGGTG	1863
Oy	1763	GACAGGGGGGCTCAGGCTGGCCGGAAGTCCCGCAAGTCAGTGTGTTACCTCCAGGGCAGT	1822
Db	1864	GACAGGGGGGCTCAGGCTGGCCGGAAGTCCCGCAAGTCAGTGTGTTACCTCCAGGGCAGT	1923
Oy	1823	GCCCTGTGTGCAACCGAGCCAGGCCCTTCCAGAGACAGAGGCGAGGGCAGGGG	1882
Db	1924	GCCCTGTGTGCAACCGAGCCAGGCCCTTCCAGAGACAGAGGCGAGGGCAGGGG	1983
Oy	1883	CAGGGAAGCCCTGTCATTCTCTACGCCAGGTTCCGAAAGTGTGTAACAGAGCCACG	1942
Db	1984	CAGGGAAGCCCTGTCATTCTCTACGCCAGGTTCCCGAAAGTGTGTAACAGAGCCACG	2043
Oy	1943	GTCGATACAGTGGAGAGGAGGGCGAGGCGTCGAGCCCTACACATAGCCACAGCTCCCTG	2002
Db	2044	GTCGATACAGTGGAGAGGAGGGCGAGGCGTCGAGCCCTACACATAGCCACAGCTCCCTG	2103
Oy	2003	ACACTGAAGAGATCCACAACCTCTTTGAGAAACACCCCTCACGTGTGTCGCGACACAT	2062
Db	2104	ACACTGAAGAGATCCACAACCTCTTTGAGAAACACCCCTCACGTGTGTCGCGACACAT	2163
Oy	2063	TTCGCTAGAGTCGGCCCAATACCCGTCACCTACAGGCTCACGCTCCGACCCGTGCTACTAG	2122
Db	2164	TTCGCTAGAGTCGGCCCAATACCCGTCACCTACAGGCTCACGCTCCGACCCGTGCTACTAG	2223
Oy	2123	GCTCACCTCCACACCCCTGTCACTACAGGCTCACTCTACAGCCTTAAGTCCAGGCGC	2182
Db	2224	GCTCACCTCCACACCCCTGTCACTACAGGCTCACTCTACAGCCTTAAGTCCAGGCGC	2283
Oy	2183	ATGTCCTCCCTGTCCAGAGGCTCAAGACTTTCTTAAGTGGATGTGAAGGACTGAAG	2242
Db	2284	ATGTCCTCCCTGTCCAGAGGCTCAAGACTTTCTTAAGTGGATGTGAAGGACTGAAG	2343
Oy	2243	TACCTTTGGGGGCAACAGCACCTTAATTTCAATCTTCACTAGCCCTGACACACTACT	2302
Db	2344	TACCTTTGGGGGCAACAGCACCTTAATTTCAATCTTCACTAGCCCTGACACACTACT	2403
Oy	2303	GTGGCAGGAATGAAAACAGAGCTTCCCGTGCAAAAAGGTCAGGCTCCACCCCGCC	2362
Db	2404	GTGGCAGGAATGAAAACAGAGCTTCCCGTGCAAAAAGGTCAGGCTCCACCCCGCC	2463
Oy	2363	CCCTCCCTGCACCTCGTCTCTCCAGTCAATCTCTGAAACACAGGCGAGGCAAG	2422
Db	2464	CCCTCCCTGCACCTCGTCTCTCCAGTCAATCTCTGAAACACAGGCGAGGCAAG	2523

	LOCUS	AX099939	2781 bp	DNA	linear	PAT 02-Apr-2001
	DEFINITION	Sequence 21 from Patent WO0120004.				
	ACCESSION	AX099939				
	VERSION	AX099939.1	GI:13538949			
	KEYWORDS					
	SOURCE	human.				
	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2781) Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R., Azizmai,V. and Lu,D.A. Protein phosphatase and kinase proteins Patent: WO 0120004-A 21-22-Mar-2001; Incyte Genomics, Inc. (US)				
	TITLE	Location/Qualifiers				
	JOURNAL	1..2781				
	FEATURES	/organism="Homo sapiens" /db_xref="taxon:9606" /note="Incyte ID NO: 5039718CB1"				
	BASE COUNT	576 a 906 c 820 g 479 t				
	ORIGIN					
Query Match	94.8%; Score 2562.8; DB 6; Length 2781;					
Best Local Similarity	96.4%; Pred. No. 0;					
Matches:2672; Conservative	0; Mismatches 2; Indels 98; Gaps 1;					
Oy	17	GCGGGTCCAGGACTGTCCGGGGGTTGAGGGAAGGGGCCGTGCCGGTGCCAGCCAGCT	76			
Db	10	GCGGGTCACGAGCTGTCCGGGGGTTGAGGGAAGGGGCCGTGCCGGTGCCAGCCAGCT	69			
Oy	77	GCTGC GGCCCTGGCTCCATTGGGCCC GTGCA CAGTAGAGC GTTGGCCCCCGGGCAGCGCG	136			
Db	70	GCTGC GGCCCTGGCTCCATTGGGCCC GTGCA CAGTAGAGC GTTGGCCCCCGGGCAGCGCG	129			
Oy	137	CCTCCAGCCCGTGGGGCC TTGGAGCACAGGGCGGTCCAGCGAAGAAGTGC ACTCCAGCGAA	196			
Db	130	CCTCCAGCCCGTGGGGCC TTGGAGCACAGGGCGGTCCAGCGAAGAAGTGC ACTCCAGCGAA	189			
Oy	197	GGCAGAGCTTTGGCGGTGCTCCGTGGGGGCTGTCTCTGGAGCTGCAGATATGAGGGGACAATG	256			
Db	190	GGCAGAGCTTTGGCGGTGCTCCGTGGGGGCTGTCTCTGGAGCTGCAGATATGAGGGGACAATG	249			
Oy	257	ATGATGCAGCAGAGGCCAGTTCTTGAGCCAACAGAAAGGCCCCGAGTGTAGAGAGAGCTCC	316			
Db	250	ATGATGCAGCAGAGGCCAGTTCTTGAGCCAACAGAAAGGCCCCGAGTGTAGAGAGAGCTCC	309			
Oy	317	ACGGGGACCACAGACACTTTCGGGCAAGGATCCACAGATCCCAAGACAGAGAGAGCAGA	376			
Db	310	ACGGGGACCACAGACACTTTCGGGCAAGGATCCCAAGATCCCAAGAGAGAGAGAGCAGA	369			

QY	377	GGCAGCACCCTGCACCTCATGTGACAGCTGCTAGAGCCCGCAGATGACATCCGCCCTGGCAG	436
Db	370	GGCACCACCTGCACCTCATGTGACAGCTGCTAGAGCCCGCAGATGACATCCGCCCTGGCAG	429
QY	437	CCGACCTGGAGCACCACCGGCGCTCCCGCGCTCCGGTACCTGGTGGAGTTTCTACACGAG	496
Db	430	CCGACCTGGAGCACCACCGGCGCTCCCGCGCTCCGGTACCTGGTGGAGTTTCTACACGAG	489
QY	497	AAGGAGAAAGTGTGAAGCCAGATGAGACGGTCTCTCGGCGCTGGAATTTCCCTGACAGCA	556
Db	490	AAGGAGAAAGTGTGAAGCCAGATGAGAGCGTCTCTCGGCGGTGATTTCCCTGACAGCA	549
QY	557	GCTCCCCAGCTGACACCTCTGGGCTGTGTTTGCCCTCTGTAGATGACACCCAGGTACT	616
Db	550	GCTCCCCAGCTGACACCTCTGGGCGCTGTGTTTGCCCTCTGTAGATGACACCCAGGTACT	609
QY	617	TGAGTGGAAACGGGGGCTTCAGCGTACGTCGTGGTGGGAAAGCCGGATCTTCAAGCCCA	676
Db	610	TGAGTGGAAACGGGGGCTTCAGCGTATGCTGTGTGGGAAAGCCGGATCTTCAAGCCCA	669
QY	677	TCTCCATCCAGACCATGTGGGGCCACACTCCAGATATTGACCAAGCATGTGAGCAGCTC	736
Db	670	TCTCCATCCAGACCATGTGGGGCCACACTCCAGATATTGACCAAGCATGTGAGCAGCTC	729
QY	737	TAGGCAAGCGGCTTTGATCCGGGTGGGAGTGTCCCTACACGTGGGCGACGACTACCAAGAG	796
Db	730	TAGGCAAGCGGCTTTGATCCGGGTGGGAGTGTCCCTACCTGGGCGACGACTACCAAGAG	789
QY	797	GACTGAACCTCCGAACAGACTGCTCCTCAATGATGATGAGCAGGCTATGGCGGCACTTGAGTCTC	856
Db	790	GACTGAACCTCCGAACAGACTGCTCCTCAATGATGATGAGCAGGCTATGGCGGCACTTGAGTCTC	849
QY	857	TGCGGCTCTCCAGCGCCGAGACCTTGGCGGGTCTCTAGAAACAGAGCAGATGAGACAGGCCA	916
Db	850	TGCGGCTCTCCAGCTGCTGCGGAGCCCTGGCGGGTCTCTAGAAACAGAGCAGATGAGACAGGCCA	909
QY	917	TCCGTCCTGAGCTGTGGAAAGTTGTGGATGTCAAGTGCCTGGAGAGTGTACTTCCAAAG	976
Db	910	TCCGTCCTGAGCTGTGGAAAGTTGTGGATGTCAAGTGCCTGGAGAGTGTACTTCCAAAG	969
QY	977	AGATCCGCGAGGCTCTGAGACCTGCGGCTGGGSGCTCCGCCCTCCAGCAAGTACCGTACTTCA	1036
Db	970	AGATCCGCGAGGCTCTGAGACCTGCGGCTGGGSGCTCCGCCCTCCAGCAAGTACCGTACTTCA	1029
QY	1037	TGAGCAACAGATGTGCTGTGCTGTGTGGACACAGCGGGACCGAGCCTTCCGCATTTTCCCC	1096
Db	1030	TGAGCAACAGATGTGCTGTGCTGTGTGGACACAGCGGGACCGAGCCTTCCGCATTTTCCCC	1089
QY	1097	ACCTTAACCTGGGCTCAGATGGAAGCAGCAAAACCTGGAGAGGCTCAGAGGAACAGGG	1156
Db	1090	ACCTTAACCTGGGCTCAGATGGAAGCAGCAAAACCTGGAGAGGCTCAGAGGAACAGGG	1149
QY	1157	TCAACCAATCTTGAACATGGCCCGGAGATTGCAACTTCTAACCTTGAGCGCTTCACT	1216
Db	1150	TCAACCAATCTTGAACATGGCCCGGAGATTGCAACTTCTAACCTTGAGCGCTTCACT	1209
QY	1217	ACCACAAATGTGCGCTCTGTGGATGAGGATGTGGCGCCACGTTGCTCGCAGCTGGAAGAGA	1276
Db	1210	ACCACAAATGTGCGCTCTGTGGATGAGGATGTGGCGCCACGTTGCTCGCAGCTGGAAGAGA	1269
QY	1277	CGCACCGCTTCATTTGAGAGCTGCAAGAGCACAGGGCACCCACGTGCTGGTCCACATGCAGAGA	1336
Db	1270	CGCACCGCTTCATTTGAGAGCTGCAAGAGCACAGGGCACCCACGTGCTGGTCCACATGCAGAGA	1329
QY	1337	TGGGCGTACAGCGGCTACAGCGCCACAGTGTGTGGGCTTATGCATATAAACAGTACAAATGCA	1396
Db	1330	TGGGCGTACAGCGGCTACAGCGCCACAGTGTGTGGGCTTATGCATATAAACAGTACAAATGCA	1389
QY	1397	GCTTGGAGCAGGCGCTTCGCGCACGTGCAAGAGCTTCGCGGCCCATTCGCGGCCCAACCTTG	1456
Db	1390	GCTTGGAGCAGGCGCTTCGCGCACGTGCAAGAGCTTCGCGGCCCATTCGCGGCCCAACCTTG	1449

Qy	1457	GCCTCTGCGCCAGCTGCGAGATCTACAGGGCATCTCTGACG-----	1497
Db	1450	GCTTCCTGCGCCAGCTGCGAGATCTACAGGGCATCTCTGACGCGCCAGCGCCAGAGCCATG	1509
Qy	1458	-----	1497
Db	1510	TCCTGGAGCAGAAAGTGGGTGGGGTCTCCCGAGAGAGCACCAGCCCTGTAAGTCTCTA	1569
Qy	1498	-----GCCAGAACCTTGAGGGTGGTGGGAGGAGAAGGTTGTAGCA	1538
Db	1570	CACCATTTCCACCTCTTCGGCCAGAAACCTGAGGGTGGTGGGAGAGAGAGTTGTAGCA	1629
Qy	1539	TGGAAAGAGAGCCAGGACAGCCCCCAAAAGAAAGAGCTTGGGACACGGCCACGTTAAAGCTCC	1598
Db	1630	TGGAAAGAGAGCCAGGACAGCCCCCAAAAGAAAGAGCTTGGGACAGGGCCACGTTAAAGCTCC	1689
Qy	1599	GAGGGGTATAGAGTTCATGATGATCTTGGAGCCCTCCCTGGAGCTGGAGAGCACTCAG	1658
Db	1690	GAGGGGTATAGAGTTCATGATGATCTTGGAGCCCTCCCTGGAGCTGGAGAGCACTCAG	1749
Qy	1659	AGACAGTGAAGATCCAGAGAGTCTTCCTCCACAGATCTTACATGAAAGGCTCTGC	1718
Db	1750	AGACAGTGAAGATCCAGAGAGTCTTCCTCCACAGATCTTACATGAAAGGCTCTGC	1809
Qy	1719	AGCCCTTCCACAGCTTGCAAGAACAGAGGAGGCGCAGAGTGGACAGGGGGGCTCAGC	1778
Db	1810	AGCCCTTCCACAGCTTGCAAGAACAGAGGAGGCGCAGAGTGGACAGGGGGGCTCAGC	1869
Qy	1779	CTGGCTTGAAGTCCCGGCGAGTCAAGTGGTTACCTCCAGGGGAGTCCCTGGTGGCCAAAC	1838
Db	1870	CTGGCTTGAAGTCCCGGCGAGTCAAGTGGTTACCTCCAGGGGAGTCCCGTGGTGGCCAAAC	1929
Qy	1839	GGAGCCAGGCTTCAGAGACAGGAGCAGGGGCGAGGGGCGAGGAGAGAGCCCTGCA	1898
Db	1930	GGAGCCAGGCTTCAGAGAGCAGGAGGCGAGGGGCGAGGGGCGAGGAGAGCCCTGCA	1989
Qy	1899	TTTCTCTACGCTCCAGGTTCCCGGAAGTGGTGGAGACAGGCGCAGCGTGAATGACAGTGAAG	1958
Db	1990	TTTTCTCTACGCTCCAGGTTCCCGGAAGTGGTGGAGACAGGCGCAGCGTGAATGAGTGAAG	2049
Qy	1959	AGGAGGGGAGGCGCTGAGCCCTCCACAGTCCGACGCTCCCTGACACTGAAGAGAGATCC	2018
Db	2050	AGGAGGGGAGGCGCTGAGCCCTCCACAGTCCGACAGTCCGACGCTCCCTGACACTGAAGAGAGATCC	2109
Qy	2019	ACAAGTCTTTGGAGAAGACACCTCAGTCTGTTGGCGCACACATTTCTCAGCTCCGC	2078
Db	2110	ACAAGTCTTTGGAGAAGACACCTCAGTCTGTTGGCGCACACATTTCTCAGCTCCGC	2169
Qy	2079	CCATACCGGTCACTACAGCTCAGCTCCGACCCCTGACACTAGGGCTACCTCCACACC	2138
Db	2170	CCATACCGGTCACTACAGCTCAGCTCCGACCCCTGACACTAGGGCTACCTCCACACC	2229
Qy	2139	CTGTCACTACAGCTCAGCTCCTACAGCTTTAAGTCCGAGGCCAGTGTGCTGTCCAA	2198
Db	2230	CTGTCACTACAGCTCAGCTCCTACAGCTTTAAGTCCGAGGCCAGTGTGCTGTCCAA	2289
Qy	2199	GGGCTCAAGACTTTTAACTGGGATGTGTAGAGGGACTGAAGTACCTTTGGGGGCAAC	2258
Db	2290	GGGCTCAAGACTTTTAACTGGGATGTGTAGAGGGACTGAAGTACCTTTGGGGGCAAC	2349
Qy	2259	AGCAGCTTAGTTTCACTTCTCAACTCTAGCCTTGACACTCACTCTGGAGAGCGGAATGAAA	2318
Db	2350	AGCAGCTTAGTTTCACTTCTCAACTCTAGCCTTGACACTCACTCTGGAGAGCGGAATGAAA	2409
Qy	2319	ACGAGACTTCCCGTCAAAAAAGGGGACAGGCTCCACACCCCGCCCTCCTGACACTCC	2378
Db	2410	ACGAGACTTCCCGTCAAAAAAGGGGACAGGCTCCACACCCCGCCCTCCTGACACTCC	2469
Qy	2379	TGTCTCTCTCCAGTTTATTTCTGTGAACAGCCAGGCGCCAGGCAACAGTGGCCCCAAAGG	2438
Db	2470	TGTCTCTCTCCAGTTTATTTCTGTGAACAGCCAGGCGCCAGGCAACAGTGGCCCCAAAGG	2529
Qy	2439	CAGGAGATCTCTCAGGCCCCAGCCGGGAGGCTTGGAAGGCTGGCAGATGGCTTCCCT	2499

DB	2530	CAGGAGGATCTCTCAGGCCCCAGCGCGGAGGCTGGAAAGGGCTGGCAGATCGTTCCT	2589
OY	2499	CATCCACCTCCACCGGTCACAGGTCTTTGCTGCTGTGCCACACCTCGTGACACACGC	2558
Db	2590	CATCCACCTCCACCGGTCAGGTCTTTGCTGCTGTGCCACACCTCGTGACACACGC	2649
OY	2559	CAGATCACAGGGCCACGAGCCAGAGATAGTCTCTTTTGTGCTTTGGGCTCGGCTCA	2618
Db	2650	CAGATCACAGGGCCACGAGCCAGAGATAGTCTCTTTTGTGCTTTGGGCTCGGCTCA	2709
OY	2619	GTCAGTTTTTCATAGCCTTACAGTATCTGCGCTTGTTACTGAGAAATTAACACATTTTCA	2678
Db	2710	GTCAGTTTTTCATAGCCTTACAGTATCTGCGCTTGTTACTGAGAAATTAACACATTTTCA	2769
OY	2679	TAAAAAATAAAAA 2690	
Db	2770	TAAAAAATAAAAA 2781	
RESULT 3			
LOCUS	AK094226		
DEFINITION	Homo sapiens cDNA FLJ36907 fis, clone BRACE2003800, weakly similar to MAP kinase phosphatase.		
ACCESSION	AK094226		
VERSION	AK094226.1 GI:21753246		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens cerebellum cDNA to mRNA, clone_1lb:BRACE2 clone:BRACE2003800.		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiro,M., Shimizu,F., Wakebe,H., Ono,T., Hishiyagi,H., Watanabe,T., Ozaki,K., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai,Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Magatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y. Nagai,K. and Isogai,T.		
TITLE	NEO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2905)		
AUTHORS	Isogai,T. and Yamamoto,J.		
TITLE	Direct Submision		
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team): 2-6-7 Kazuo-Kamatari, Kisanazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:01-438-52-3973, Fax:01-438-52-3986)		
COMMENT	NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; Clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.		
FEATURES	Location/Qualifiers		
SOURCE	1: .2905		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="BRACE2003800"		
	/tissue_type="cerebellum"		
	/clone_1lb="BRACE2"		
	/note="Cloning vector: pME18SFL3"		
	646..2187		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAAC04314.1"		
	/db_xref="GI:21753247"		
	/translation="MAPPLSPASSPSCGLGLPLWMSDTQVYLDGSGFSVTSGGQSRIFKRSISQTMWATLQVHQAQEAALVGSGLVPGGSALTWASHYQDRLSQSLNEMWRA		

BASE COUNT	591 a	940 c	867 g	507 t
ORIGIN				
Query Match	87.0%	Score 2352, 8;	DB 9;	Length 2905;
Best Local Similarity	91.3%;	Pred. No. 0;		
Matches 2644;	Conservative	0;	Mismatches 2;	Indels 250;
				Gaps 4;
1	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	90		
2	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	60		
3	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	150		
4	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	120		
5	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	158		
6	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	180		
7	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	201		
8	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	236		
9	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	261		
10	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	296		
11	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	321		
12	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	356		
13	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	381		
14	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	416		
15	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	441		
16	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	476		
17	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	501		
18	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	536		
19	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	561		
20	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	596		
21	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	621		
22	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	656		
23	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	681		
24	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	716		
25	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	741		
26	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	776		
27	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	801		
28	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	836		
29	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	861		
30	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	886		
31	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	911		
32	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	936		
33	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	961		
34	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	986		
35	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1011		
36	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1036		
37	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1061		
38	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1086		
39	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1111		
40	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1136		
41	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1161		
42	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1186		
43	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1211		
44	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1236		
45	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1261		
46	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1286		
47	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1311		
48	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1336		
49	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1361		
50	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1386		
51	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1411		
52	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1436		
53	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1461		
54	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1486		
55	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1511		
56	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1536		
57	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC	1561		
58	GTCCGCGGGGTTGAGAGGAAAGGGCCGTCGCCAGCCAGCCAGGTGCTCGCGCCCTGGC			

Dh 837 ATGTGAGGACGCTTAGAGCAGCGGCTTTGTACCGGGGTGGAGTGCCTTACCTTGAGGCGAC 896
Oy 783 CCACCTACGAGAGAGAGACTGAACTCCGAACAGAGAGTGCCTCAATGATGAGGAGGCTATGGC 842
Dh 897 CCACCTACGAGAGAGAGACTGAACTCCGAACAGAGAGTGCCTCAATGATGAGGAGGCTATGGC 956
Oy 843 CGACCTGAGAGTCTGCGGCGCTCCAGCGCGAGAGCTGGCGGGTCTCTAGAAACAGAGACA 902
Dh 957 CGACCTGAGAGTCTGCGGCGCTCCAGCGCGAGAGCTGGCGGGTCTCTAGAAACAGAGACA 1016
Oy 903 GATGAGAGAGGAGATCGGTGCTGAGCTGTGGAAGTGTGGATGTGAGTGCCTGAGAG 962
Dh 1017 GATGAGAGAGGAGATCGGTGCTGAGCTGTGGAAGTGTGGATGTGAGTGCCTGAGAG 1076
Oy 963 TGTCACTTCCAAAGAGATCCGCGAGGCTCTGGAGCTGGCGGCTGGGGGCTCCCTCCAGCA 1022
Dh 1077 TGTCACTTCCAAAGAGATCCGCGAGGCTCTGGAGCTGGCGGCTGGGGGCTCCCTCCAGCA 1136
Oy 1023 GTACCGTGACTTCATGACACCAAGATGCTGCTGCTGTGTCACAGCGGAGCCGAGCTC 1082
Dh 1137 GTACCGTGACTTCATGACACCAAGATGCTGCTGCTGTGTCACAGCGGAGCCGAGCTC 1196
Oy 1083 CCGCATCTTCCCCCACTTACCTTGGGCTCAGAGTGGAAACGAGCAAACTGGAGAGCT 1142
Dh 1197 CCGCATCTTCCCCCACTTACCTTGGGCTCAGAGTGGAAACGAGCAAACTGGAGAGCT 1256
Oy 1143 GCAGAGAGAACAGGAGTACCCACCATCTTGAACATGGCCCGGAGATTGACAACTTGTACCC 1202
Dh 1257 GCAGAGAGAACAGGAGTACCCACCATCTTGAACATGGCCCGGAGATTGACAACTTGTACCC 1316
Oy 1203 TTAGGGCTTCACTTACCAATATGTGGCTCTGGGATGAGGAGTGGCCGACACTGTGCG 1262
Dh 1317 TTAGGGCTTCACTTACCAATATGTGGCTCTGGGATGAGGAGTGGCCGACACTGTGCG 1376
Oy 1263 GCATGGAAGAGAGAGCAACGCTTATTGAGGCTGCAAGAGACAGAGGAGCCAGCTGCT 1322
Dh 1377 GCATGGAAGAGAGAGCAACGCTTATTGAGGCTGCAAGAGACAGAGGAGCCAGCTGCT 1436
Oy 1323 GGTCCACTGCAAGATGGGCGTCAAGCGGCTCAGCGGCGACAGTGTCTGATCCATGAA 1382
Dh 1437 GGTCCACTGCAAGATGGGCGTCAAGCGGCTCAGCGGCGACAGTGTCTGATCCATGAA 1496
Oy 1383 GCAGTACGAATGACGCTTGAGAGAGGCTGGCGCACTGTCAGAGAGTCCGGCCCATGCC 1442
Dh 1497 GCAGTACGAATGACGCTTGAGAGAGGCTGGCGCACTGTCAGAGAGTCCGGCCCATGCC 1556
Oy 1443 CGGCCCAACCGTTGGGCTTCTGCGAGCTGCAAGTCTACAGGGGATCTGTAGC----- 1497
Dh 1557 CGGCCCAACCGTTGGGCTTCTGCGAGCTGCAAGTCTACAGGGGATCTGTAGC----- 1616
Oy 1498 ----- 1497
Dh 1617 CGGCCAGACCATGTCTGGGAGCAGAAAGTGGGTGGGCTCTCCCAAGAGAGACCCACAC 1676
Oy 1498 -----GCCAGAACTGAGGGTGGTGGGAGCA 1524
Dh 1677 CCCTGAAGTCTCTAGACCAATTCACCTTTCGCCAGAACTGAGGGTGGTGGGAGCA 1736
Oy 1525 GAAGGTTGTAAGCATGGAAGAGAGAGCCAGGCGCCGGAAGAAAGAGCTGGGCGACAGGCC 1584
Dh 1737 GAAGGTTGTAAGCATGGAAGAGAGAGCCAGGCGCCGGAAGAAAGAGCTGGGCGACAGGCC 1796
Oy 1585 ACGTATAACTCCGAGGGGTGATGAGTCCATCAATCTTCTGAGAGCCCTCTTGGAGCT 1644
Dh 1797 ACGTATAACTCCGAGGGGTGATGAGTCCATCAATCTTCTGAGAGCCCTCTTGGAGCT 1856
Oy 1645 GAGAGACACCTCAGAGACCAAGTACATGCGCAAGAGCTTCTCTCCACGAGTCTTACCA 1704
Dh 1857 GAGAGACACCTCAGAGACCAAGTACATGCGCAAGAGCTTCTCTCCACGAGTCTTACCA 1916
Oy 1705 TGAAGAGCCTTCGACAGCCCTTCCCAAGCTTGCAGAGACCAAGGAGGCGCAGAGGTGGA 1764
Dh 1917 TGAAGAGCCTTCGACAGCCCTTCCCAAGCTTGCAGAGACCAAGGAGGCGCAGAGGTGGA 1976

Oy 1765 CAGGGGGCCTCAGCCTTGCCCTGAAAGTCCCGCACTGACTGTTACCTCCAGGGCAGTGC 1824
Dh 1977 CAGGGGGCCTCAGCCTTGCCCTGAAAGTCCCGCACTGACTGTTACCTCCAGGGCAGTGC 2036
Oy 1825 CGTGGTGGCCAAACCGGAGCCGAGGCTTCCAGAGCAGAGAGAGGAGGAGGAGGAGCA 1884
Dh 2037 CGTGGTGGCCAAACCGGAGCCGAGGCTTCCAGAGCAGAGAGAGGAGGAGGAGGAGCA 2096
Oy 1885 GGGAGAGCCCTGCATTTCCTCTACGCCAGGTTCCGGAAGTGTGAGACAGGCGAGCGT 1944
Dh 2097 GGGAGAGCCCTGCATTTCCTCTACGCCAGGTTCCGGAAGTGTGAGACAGGCGAGCGT 2156
Oy 1945 GCATGACAGTGAAGAGAGAGGCGAGGCTGAGCCTTCAACATGCGCCAGCTCCCTTAC 2004
Dh 2157 GCATGACAGTGAAGAGAGGCGAGGCTGAGCCTTCAACATGCGCCAGCTCCCTTAC 2216
Oy 2005 ACTGAAGAGATCCCACTCTCTTGGAAACACCTTCAGCTGTGTGCGGACACATTC 2064
Dh 2217 ACTGAAGAGATCCCACTCTCTTGGAAACACCTTCAGCTGTGTGCGGACACATTC 2276
Oy 2065 CTCTCAGCTCCGCCCATACCGGTCTACAGCCTACCTCCACCCCTGTCTACAGGC 2124
Dh 2277 CTCTCAGCTCCGCCCATACCGGTCTACAGCCTACCTCCACCCCTGTCTACAGGC 2336
Oy 2125 CTCACCTCCACCCCTGTCTACAGCCTACCTCTTACAGCTTAAAGTCCAGGCGCAT 2184
Dh 2337 CTCACCTCCACCCCTGTCTACAGCCTACCTCTTACAGCTTAAAGTCCAGGCGCAT 2396
Oy 2185 GTCTGCTGTCCCAAGGGCTCAAGACTTCTTAACTGGAGTGTGTAGAGGAGTGAAGGTA 2244
Dh 2397 GTCTGCTGTCCCAAGGGCTCAAGACTTCTTAACTGGAGTGTGTAGAGGAGTGAAGGTA 2456
Oy 2245 CCTTTGGGGGCAACAGACCCCTAGTTTCATTCATCTACCTGACACTGACTGT 2304
Dh 2457 CCTTTGGGGGCAACAGACCCCTAGTTTCATTCATCTACCTGACACTGACTGT 2516
Oy 2305 GGCACGGGAATGAACACAGAGCTTCCCGTGCAGAAAGGGTCAACGGCTCCACCCCGGCC 2364
Dh 2517 GGCACGGGAATGAACACAGAGCTTCCCGTGCAGAAAGGGTCAACGGCTCCACCCCGGCC 2576
Oy 2365 CTCCGTCACGCTGTCTGCTCTCCAGTTTCATTCCTGGAACAGGAGGAGGCAACCA 2424
Dh 2577 CTCCGTCACGCTGTCTGCTCTCCAGTTTCATTCCTGGAACAGGAGGAGGCAACCA 2636
Oy 2425 GTGGCCCCCAAGGAGGAGGATCCTCAGGCCCAAGCCGCGGAGGCTGGAAGGGCTGG 2484
Dh 2637 GTGGCCCCCAAGGAGGAGGATCCTCAGGCCCAAGCCGCGGAGGCTGGAAGGGCTGG 2696
Oy 2485 CAGATCGCTTCCCTCATCCACTCCACGGGTCCAGGTCTTGTCTGTCCCAACCTC 2544
Dh 2697 CAGATCGCTTCCCTCATCCACTCCACGGGTCCAGGTCTTGTCTGTCCCAACCTC 2756
Oy 2545 CTGTGACACAGGCGCAAGATCAGAGGCAACAGGAGATGTCTGTTTGTCTT 2604
Dh 2757 CTGTGACACAGGCGCAAGATCAGAGGCAACAGGAGATGTCTGTTTGTCTT 2816
Oy 2605 CTGGCCTCTGGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGGCTTGTACTGGAAT 2664
Dh 2817 CTGGCCTCTGGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGGCTTGTACTGGAAT 2876
Oy 2665 AAAACACATTTTCATA 2680
Dh 2877 AAAACACATTTTCATA 2892

RESULT 4
AA451343
LOCUS AA451343 1980 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent W00224740.
ACCESSION AX451343
VERSION AX451343.1 GI:21698394
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1
AUTHORS Luche,R.M. and Wei,B.
TITLE Dsp-15 dual-specificity phosphatase
JOURNAL Patent: WO 0224740-A 1 28-MAR-2002;
Ceptyr, Inc. (US)
FEATURES
Source Location/Qualifiers
1..1980
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 409 a 616 c 640 g 315 t
ORIGIN
Query Match 65.6%; Score 1774; DB 6; Length 1980;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;
QY 94 ATGGCCCTGGCTCAGTACAGTGGCGCTTCGCGGACAGCGCGCTCCACAGCGCGTGGGG 153
Db 1 ATGGCCCTGGCTCAGTACAGTGGCGCTTCGCGGACAGCGCGCTCCACAGCGCGTGGGG 60
QY 154 CCCTGGACACAGCGCGCTCCAGCGAGAGTGCATCCAGCGAGGCGAGAGCTTTGCGGTG 213
Db 61 CCCTGGACACAGCGCGCTCCAGCGAGAGTGCATCCAGCGAGGCGAGAGCTTTGCGGTG 120
QY 214 CTCCTGGGCGCTGCTCTGGGACCTGAGATGGAGGGGACATGATGATGACAGAGGCC 273
Db 121 CTCCTGGGCGCTGCTCTGGGACCTGAGATGGAGGGGACATGATGATGACAGAGGCC 180
QY 274 AGTTGTAGCCAAAGAGAGAGGCGCGGAGTGAAGAGAGCTCCAGGGGACAGACAGAC 333
Db 181 AGTTGTAGCCAAAGAGAGAGGCGCGGAGTGAAGAGAGCTCCAGGGGACAGACAGAC 240
QY 334 TTCGGGCAAGGATCCCAAGATGCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGCTG 393
Db 241 TTCGGGCAAGGATCCCAAGATGCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGCTG 300
QY 394 ATGTGACAGCTGCTGAGGCGCGAGATGACATCCGCTGGGACGCCAGCTGGAGGACCC 453
Db 301 ATGTGACAGCTGCTGAGGCGCGAGATGACATCCGCTGGGACGCCAGCTGGAGGACCC 360
QY 454 CGGCTTCCCGGCTCCGCTGACCTGCTGATTTCTACACGAGAGAGAGAGAGTGTGAGC 513
Db 361 CGGCTTCCCGGCTCCGCTGACCTGCTGATTTCTACACGAGAGAGAGAGAGTGTGAGC 420
QY 514 CAGATGAGAGCGGCTCTCTGGGCGTGGATTTCCTGACAGAGCTCCCGCAGCTGCACC 573
Db 421 CAGATGAGAGCGGCTCTCTGGGCGTGGATTTCCTGACAGAGCTCCCGCAGCTGCACC 480
QY 574 CTGGGCTGCTGCTTCCCTCTGGAGTGACACCCAGGTGTACTTATGATGAGAGCGGGGC 633
Db 481 CTGGGCTGCTGCTTCCCTCTGGAGTGACACCCAGGTGTACTTATGATGAGAGCGGGGC 540
QY 634 TTCACGTCGAGCTGCTGGGCGAAAGCGGAGTCTCAAGCCCATCTCATCCAGACATG 693
Db 541 TTCACGTCGAGCTGCTGGGCGAAAGCGGAGTCTCAAGCCCATCTCATCCAGACATG 600
QY 694 TGGGCGACACTCCAGTATTGACCAACAGATGTAGGAGCTCTAGGAGGCGCTTGTGA 753
Db 601 TGGGCGACACTCCAGTATTGACCAACAGATGTAGGAGCTCTAGGAGGCGCTTGTGA 660
QY 754 CCGGCTGCACTGCTCTACCTGCGGCGACCACTACAGAGAGAGACTAACTCCGAGAC 813
Db 661 CCGGCTGCACTGCTCTACCTGCGGCGACCACTACAGAGAGAGACTAACTCCGAGAC 720
QY 814 AGCTCCCTCAATGATGTAGAGGCTATGGCGACCTGAGTCTCTGCGGCTCCCGCGCC 873
Db 721 AGCTCCCTCAATGATGTAGAGGCTATGGCGACCTGAGTCTCTGCGGCTCCCGCGCC 780
QY 874 GAGCCTGGCGGGTCTCTCAGAACAGAGAGAGATGAGAGGCGATCCGTGAGCTGTGG 933

Db 781 GAGCCTGGCGGGTCTCTCAGAACAGAGAGAGATGAGAGGCGATCCGTGAGCTGTGG 840
QY 934 AAAGTGTGATGTAGTACCTGAGAGTGTCACTTCCAAAGAGATCCCGAGGCTCTG 993
Db 841 AAAGTGTGATGTAGTACCTGAGAGTGTCACTTCCAAAGAGATCCCGAGGCTCTG 900
QY 994 GAGCTGGCGCTGGGCTCCCTCCAGAGTACGCTGAGTACCTGATGACACACAGATGCTG 1053
Db 901 GAGCTGGCGCTGGGCTCCCTCCAGAGTACGCTGAGTACCTGATGACACACAGATGCTG 960
QY 1054 CTGCTGTGGGACACAGCGGAGGAGGAGCTCCGCACTCTTCCCGACCTGTACCTGGGCTCA 1113
Db 961 CTGCTGTGGGACACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
QY 1114 GAGTGAACGACGCAAACTGAGAGAGTGCAGAGAGAGAGAGGAGGAGGAGGAGGAGG 1173
Db 1021 GAGTGAACGACGCAAACTGAGAGAGTGCAGAGAGAGAGAGGAGGAGGAGGAGGAGG 1080
QY 1174 ATGGCCGCGGAGATTGACACCTTACCTGAGGCTTACCTGACATGATGAGGCGCTC 1233
Db 1081 ATGGCCGCGGAGATTGACACCTTACCTGAGGCTTACCTGACATGATGAGGCGCTC 1140
QY 1234 TGGATGAGAGTCCGCGCAGCTGCTGCGCACTGGAAGAGAGAGAGAGAGGAGGAGG 1293
Db 1141 TGGATGAGAGTCCGCGCAGCTGCTGCGCACTGGAAGAGAGAGAGAGGAGGAGGAGG 1200
QY 1294 GCTGCAAGAGCAGAGGAGCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1353
Db 1201 GCTGCAAGAGCAGAGGAGCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1354 GCGGCGACAGTGTGGGCTATGGCTGGAAGAGTGCATGATGATGATGATGATGATGATG 1413
Db 1261 GCGGCGACAGTGTGGGCTATGGCTGGAAGAGTGCATGATGATGATGATGATGATGATG 1320
QY 1414 CGCCAGCTGACAGAGAGCTCCGCGGCTATGCGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1473
Db 1321 CGCCAGCTGACAGAGAGCTCCGCGGCTATGCGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1380
QY 1474 CAGATCTACAGAGGAGTCTGACG----- 1497
Db 1381 CAGATCTACAGAGGAGTCTGACG----- 1440
QY 1498 ----- 1497
Db 1441 GGTGGGCTCTCCCGAGAGAGAGCAGCCAGCCCTGTAAGTCTTACACCTTCCACCTCTT 1500
QY 1498 --GCCAGAACTGAGGAGTGTGGGAGAGAGAGTGTGATGAGCATGGAAGAGAGCCAGCA 1555
Db 1501 CGGCAGAACTGAGGAGTGTGGGAGAGAGAGTGTGATGAGCATGGAAGAGAGCCAGCA 1560
QY 1556 GCGCCGAAAGAAAGAGCTGCGGCGACAGGCGACGATATAAAGCTCCGAGGGGCTATGAGTCC 1615
Db 1561 GCGCCGAAAGAAAGAGCTGCGGCGACAGGCGACGATATAAAGCTCCGAGGGGCTATGAGTCC 1620
QY 1616 ATCAGTCTTTCGAGGAGCTCCCTTGGAGCTGGAGACCACTGAGAGACAGAGACATGCA 1675
Db 1621 ATCAGTCTTTCGAGGAGCTCCCTTGGAGCTGGAGACCACTGAGAGACAGAGACATGCA 1680
QY 1676 GAGGCTTCTCTTCCACAGAGTCTTACATGAAGAGCTCTGACAGCCCTTCCACAGCTT 1735
Db 1681 GAGGCTTCTCTTCCACAGAGTCTTACATGAAGAGCTCTGACAGCCCTTCCACAGCTT 1740
QY 1736 GCAAGAGCAAAAGAGGCGCAGAGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1795
Db 1741 GCAAGAGCAAAAGAGGCGCAGAGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
QY 1796 CAGTCAGTGTATCCCTCAGGAGGAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1855
Db 1801 CAGTCAGTGTATCCCTCAGGAGGAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
QY 1856 GAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1915

Db	1861	GAGCAGAGACGAGGGCGAGGGGCGAGGGGCGACGGGAGGAGAGGCCCTGCATTTCTCTACAGCCCAAG	1920
Qy	1916	TTCCGGAAGGTGTGTGAGACAGAGCCAGCGCTGCATGACAGTGGAGAGAGGGCGAGGCGCTGA	1975
Db	1921	TTCCGGAAGGTGTGTGAGACAGAGCCAGCGCTGCATGACAGTGGAGAGAGGGCGAGGCGCTGA	1980
RESULT 5			
AX398774		2322 bp	DNA
LOCUS			Linear
DEFINITION	Sequence 10 from Patent W00220732.		PAT 27-May-2002
ACCESSION	AX398774		
VERSION	AX398774.1	GI:21261307	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Liou,J.R.		
TITLE	Regulation of human map kinase phosphatase-like enzyme		
JOURNAL	Patent: WO 0220732-A 10 14-MAR-2002;		
	Bayer Aktiengesellschaft (DE)		
FEATURES	Location/Qualifiers		
Source	1..2322		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	467 a 751 c 718 g 386 t		
ORIGIN			
Query Match	62.8%; Score 1699; DB 6; Length 2322;		
Best Local Similarity	90.9%; Pred. No. 0;		
Matches 1987; Conservative	0; Mismatches 75; Indels 123; Gaps 12.		
Qy	132	CGGCGCTCCAGCCCGCGGGCCCGGGAGCAGGCGGCTCCAGCGAAGAGTCCGACTCCA	191
Db	147	CGCCCTCACCCCTGGGGCTCTCTCTGCGAGAGACTGTGCGAGGAGTGCAGTCTCA	206
Qy	192	GCGAAGCGAGACCTTTCGCGTCTCC-GTGGGCGTGTCTCTGGAGTGCAGATGGAGGG	250
Db	207	GCGA----AGAGCTTGGCGGTCTCTCTGTGGGGCTGTCTCTGGGACTGCAGATGGAGGGG	262
Qy	251	ACAATGATGATGCAGCAGAGGCCAGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAAGG	310
Db	263	ACAATGATGATGCAGCAGAGGCCAGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAAGG	322
Qy	311	AGCTCCAGGGG---ACGAGACAGATT--CGGGCAGAGATCCAGAGTCCCGAGAGCA	365
Db	323	AGCTCCAGGGGAGACCCAGACAGACTTTCGTTGCAAGAGATCCAGAGTCCCGAGAGCA	382
Qy	366	GGAGGAGCAGAGGACACCTGCACCTCATGTGTACAGCTTGAGGCCCGCAGAGATGACAT	425
Db	383	GGAGGAGCAGAGGACACCACTCATGTGTACAGCTTGAGGCCCGCAGAGATGACAT	442
Qy	426	CGCGCTTGCA--GCCAGCTGAGAGCACCCTCCCGGCTCCCGCTACCTGCGCTA---478	
Db	443	CGGAGTGGAGAGCCCAAGCTGGAGGACCCCGGGGCTCCCGGGATCCGATACCTTGC	502
Qy	479	TGCTATTCTACAGAGAG--GAGAGGCTGAG--CGAGATGATAGAGGTCTCTCTGGG	536
Db	503	TGCTATTCTACAGAGAGGAGAGAGGTCTGAGCCAGAGATGATAGAGGTCTCTCTGGG	562
Qy	537	-CGTGATTTCCTGCACAGAGTCCCGCAGCTGCAACCCCTGGGCTGTCTTCCCGCTCT	595
Db	563	ACGTGATTTCCCTGTACAGAGTCCCGCAGCTGCAACCCCTGGGCTGTCTTCCCGCTCT	622
Qy	596	GGAAGCAGACCCAGGTGTACTTATGAGAGACGGGGCTTTCAGCGTGTGTGGGCT	655
Db	623	GGAGTACACCCAGGTGTACTTATGAGAGACGGGGCTTTCAGCGTGTGTGGGCT	682
Qy	656	AAAGCGGATCTTCAAGCCCATCTCATGCAGACATGTGTGGCCACACTTCAGATATTGC	715
Db	683	AAAGCGGATCTTCAAGCCCATCTCATGCAGACATGTGTGGCCACACTTCAGATATTGC	742

OY	716	ACCAAGCATGTGAGGACAGCTCTTAGCGAGCGGCTTGTACCGGGTGGCAGTGCCTCAACT	775
Db	743	ACCAAGCATGTGAGGAGAGCTTAGGCGAGGGGCTTGTACCGGGTGGCAGTGCCTCAACT	802
OY	776	GGGCGACCACTACCGAGGAGAGCTCAATCCGAGACAGAGCTCAATGAGTGGACGG	835
Db	803	GGGCGACCACTACCGAGGAGAGCTCAATCCGAGAGAGCTGCTCAATGAGTGGACGG	862
OY	836	CTATGGCGGACCTGGAGTCTCTCGGCGCTCCAGCGCGGACCTGGCGGCTCTCAGAAC	895
Db	863	CTATGGCGGACCTGGAGTCTCTCGGCGCTCCAGCGCGGACCTGGCGGCTCTCAGAAC	922
OY	896	AGGAGCAGATGAGAGCAGCGGCAATCCGGTGCAGACTGTGGAAAGTGTGGATGTCAGTGAC	955
Db	923	AGGAGCAGATGAGAGCGGGGCATCCGGTGCAGCTGGAAAGTGTGGATGTCAGTGAC	982
OY	956	TGAGAGATGTACTTCCAAAGAGATCCGGCAGGCTCTGGAGTGTGCGCTGGGGCTCCCC	1015
Db	983	TGAGAGATGTGCGTCTCCAGAGATCCGGCAGGCTCTGGAGTGTGCGCTGGGGCTCCCC	1042
OY	1016	TCACAGCAGTACCGTGACTTATGACAAACCAATGCTGTGCTGGTGGCACAAGGGGAC	1075
Db	1043	TCACAGCAGTACCGTGACTTATGACAAACCAATGCTGTGCTGGTGGCACAAGGGGAC	1102
OY	1076	GAGCCTCCCGCATTTTCCCCACCTCTACCTTGAGGCTCAGAGTGAACACAGCAAACTGG	1135
Db	1103	GAGCCTCCCGCATTTTCCCCACCTCTACCTTGAGGCTCAGAGTGAACACAGCAAACTGG	1162
OY	1136	AGGACCTGCAGAGACAGGATCACCCACATCTTGAACATGAGCCGGGAGATTGACAACT	1195
Db	1163	AGGACCTGCAGAGAGAACAGGGTACCACCACTTTG--ATGGCCGGGAGATTGACAACT	1219
OY	1196	TCTACCCGAGCGGTTTACCTTACCACAAATGTGGCGCTTGSGATGAGAGATCGGCCAGC	1255
Db	1220	TCTACCCGAGCGGTTTACCTTACCACAAATGTGGCGCTTGSGATGAGAGATCGGCCAGC	1279
OY	1256	TGCTGCGCACTGGAGAGAGACGACCGCTTATTGAGGCTGCAAAGACACAGGCAACC	1315
Db	1280	TGCTGCGCACTGGAGAGAGACGACCGCTTATTGAGGCTGCAAAGACACAGGCAACC	1339
OY	1316	ACGTGCTGTGCTACCTGCAAGATGGGGGTACCGCGTCACGGGCAACAGTGGGGCTATG	1375
Db	1340	ACGTGCTGTGCTACCTGCAAGATGGGGGTACCGCGTCACAGGCAAGTGGGGCTATG	1399
OY	1376	CCATGAACAGTACGAATGCAAGCTGTGAGCAGGCGCTTGGCGCAGCTGCAGAGCTCGCGC	1435
Db	1400	CCATGAACAGTACGAATGCAAGCTGTGAGCAGGCGCTTGGCGCAGCTGCAGAGCTCGCGC	1459
OY	1436	CCATGCGCGGCGCCCAACCTGTGCTCTGGCGCAGTGTGAGATACCAAGGGCATCTCGA	1495
Db	1460	CCATGCGCGGCGCCCAACCTGTGCTCTGGCGCAGTGTGAGATACCAAGGGCATCTCGA	1519
OY	1496	CG-----CG-----	1497
Db	1520	CGGCGCAGCGCGCAGAGGCATGTCTGGGAGCAGAAATGGGGTCTCCCAAGAGAGC	1579
OY	1498	-----GCCAAGACTGAGGTTGGT	1517
Db	1580	ACCCAGCCCTGAAGTCTCTACACCATTTCCACCTCTTCCGCAAGAACTGAGGGTGGT	1639
OY	1518	GGGAGGAGAAAGTTGTAGGCAATGGAAACAGAGCGACCGCCGAAAGAAAGAGCTGTGGC	1577
Db	1640	GGGAGGAGAAAGTTGTAGGCAATGGAAACAGAGCGACCGCCGAAAGAAAGAGCTGTGGC	1699
OY	1578	CACGGCCACGTAATAAACCTCCGAGGGGTCTGAAGTTCATCACTTCTTGAAGCCCTCT	1637
Db	1700	CACGGCCACGTAATAAACCTCCGAGGGGTCTGAAGTTCATCACTTCTTGAAGCCCTCT	1759
OY	1638	TGGAGCTGGAAGAGCACTCAGAGACAGCTGACATGCCAAGAGTCTTCTTCCACAGAGT	1697
Db	1760	TGGAGCTGGAAGAGCACTCAGAGACAGCTGACATGCCAAGAGTCTTCTTCCACAGAGT	1819

QY	1698	CTACATGAGAGAGCTCTGACAGCCCTTCCACACACTTTGACAAAGACCAAGGAGGCGACG	1757
Db	1820	CTTCACATGAAAGAGCTCTGACAGCCCTTCCACACACTTTGACAAAGGACCAAGGAGGCGACG	1879
QY	1758	AGGTGGACAGGGGGGCTCAGCCTCGCCCTGAAGTCCCGGACGTCACTGTGTACCTTCACAG	1817
Db	1880	AGGTGGACAGGGGGGCTCAGCCTCGCCCTGAAGTCCCGGACGTCACTGTGTACCTTCACAG	1939
QY	1818	GCAGTGGCGGTGTGGCCCAACGGACCCAGCGCCTTCACAGAGCAGGAGCAAGGGGACAGGGGC	1877
Db	1940	GCAGTGGCGGTGTGGCCCAACGGACCCAGCGCCTTCACAGAGCAGGAGCAAGGGGACAGGGGC	1999
QY	1878	AGGGGACAGGAGAGCCCTGACATTTCTCTTACGCCCCAGGTTCCGGAAGGTGAGAACAGG	1937
Db	2000	AGGGGACAGGAGAGCCCTGACATTTCTCTTACGCCCCAGGTTCCGGAAGGTGAGAACAGG	2059
QY	1938	CCAGGCTGCATGACAGTGGAGAGAGAGAGGGGCGAGGCTGAGCCTTCACACATGCCCCAGCTC	1997
Db	2060	CCAGGCTGCATGACAGTGGAGAGAGAGGGGCG - GGCTTGAGCCCTCACACATGCCCCAGCTC	2117
QY	1998	CCCTCACACTGAGAGAGATCCACACACTCTTTGAGAAACACCCTCAGCTCTGTTCGCCCA	2057
Db	2118	CCCTCACACTGAGAGAGATCCACACACTCTTTGAGAAACACCCTCAGCTCTGTTCGCCCA	2177
QY	2058	CACATTCCTCTCAGTCTCGGCCCATATCCCGTCACTACAGGCTCACTCTCCACCCCTGTCA	2117
Db	2178	CACATTCCTCTCAGTCTCGGCCCATATCCCGTCACTACAGGCTCACTCTCCACCCCTGTCA	2237
QY	2118	CTACGGCTTCACACTCCACCCCTGTCACTACAGGCTCAGCTCACTCTCAAGCCTTAACTCCA	2177
Db	2238	CTACGGCTTCACACTCCACCCCTGTCACTACAGGCTCAGCTCACTCTCAAGCCTTAACTCCA	2297
QY	2178	GGCCCATGTCTGCTCTGTCCAAAGGC 2202	
Db	2298	GGCCCATGTCTGCTCTGTCCAAAGGC 2222	

RESULT 6	
BC004176	
LOCUS	BC004176 1905 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, similar to hypothetical protein FLJ10928, clone MGC:2772 IMAGE:2958967, mRNA, complete cds.
ACCESSION	BC004176
VERSION	BC004176.1 GI:13278815
KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1905)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	
Contact: MGC help desk	
Email: cgabds-rt@mail.nih.gov	
Tissue Procurement: DCTD/DRP	
CDNA Library Preparation: Rubin Laboratory	
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
DNA Sequencing by: Institute for Systems Biology	
http://www.systemsbiology.org	
contact: amadansystemsbiology.org	
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan	
<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov</p> <p>Series: IRAL Plate: 2 Row: m Column: 8.</p> <p>Location/Qualifiers</p> <p>1..1905</p>	
FEATURES	
source	

Query Match	Best Local Similarity	Matches 1789;	Conservative	Mismatches	Indels	Gaps
61.6%; Score 1665.2;	94.4%; Pred. No. 0;	0;	8;	99;	2;	
DB 9;	Length 1905;					
Query 907	GAGCGAGCGATCCGTCGTCAGCTGTGGAAAGTGTGATGTCAGTACGACTGGAGCTCTC	64				
DB 6	GAGGAGGCGATCCGTGCT-AGCTGTGGAAAGTGTGATGTCAGTACGACTGGAGCTGTC	64				
Query 967	ACTTCCAAAGAGATCCGCGAGGCTCTGGAGCTGCGCCCTGGGGCTCCCTCCAGCAGTAC	1026				
DB 65	ACTTCCAAAGAGATCCGCGAGGCTCTGGAGCTGCGCCCTGGGGCTCCCTCCAGCAGTAC	124				
Query 1027	CCTGACTTCATGACAAACAGATGCTGCTGCTGTGTGACACAGCGGACCGAGCTCCGCC	1086				
DB 125	CCTGACTTCATGACAAACAGATGCTGCTGCTGTGTGACACAGCGGACCGAGCTCCGCC	184				
Query 1087	ATCTTCCCCACCTCTACTCTGGGCTCAGAGTGGAAAGCGAGCAACCTTGAGAGCTGCAG	1146				
DB 185	ATCTTCCCCACCTCTACTCTGGGCTCAGAGTGGAAAGCGAGCAACCTTGAGAGCTGCAG	244				
Query 1147	AGGAACAGGGGTACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCAACCTTGAG	1206				
DB 245	AGGAACAGGGGTACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCAACCTTGAG	304				
Query 1207	GCCTTCACCTACCAAAATGTGGCGCTCTTGGATGTAGAGAGTGGGCCACTGTGCGCGAC	1266				
DB 305	GCCTTCACCTACCAAAATGTGTGGCTCTTGGATGTAGAGAGTGGGCCACTGTGCGCGAC	364				
Query 1267	TGGAAAGAGACACACCGCTCATTTAGAGCTGCAABAGACACAGGGACCCAGCTGTGTC	1326				
DB 365	TGGAAAGAGACACACCGCTCATTTAGAGCTGCAABAGACACAGGGACCCAGCTGTGTC	424				
Query 1327	CACCTGCAAGATGGGCGTACAGCGCTCAGGGCGCACAGTGTGGCTATGCCATGAAGCAG	1386				
DB 425	CACCTGCAAGATGGGCGTACAGCGCTCAGGGCGCACAGTGTGGCTATGCCATGAAGCAG	484				
Query 1387	TTCGAAATGAGCTGTGAGAGAGGCGCTTGGCGCACGTGACAGAGACTCCGGCCCATGCGCCG	1446				
DB 485	TTCGAAATGAGCTGTGAGAGAGGCGCTTGGCGCACGTGACAGAGACTCCGGCCCATGCGCCG	544				
Query 1447	CCCAACCCCTGGCTTCTGTGCGCAGCTGAGATCTACAGAGGGCATCTACAG-----	1497				
DB 545	CCCAACCCCTGGCTTCTGTGCGCAGCTGAGATCTACAGAGGGCATCTACAGCGCACGCGC	604				
Query 1498	-----	1497				
DB 605	CAGAGCATGTCTGGGAGCAGAAAGTGGGTGGGCTCTCCACAGAGAGACACCCAGCCCT	664				
Query 1498	-----GCCAGAACCTGAGGGGTGTGGGAGAGAGAG	1528				
DB 665	GAAAGTCTACACCAATTCCCACTCTTCCGCCAGAACTGAGAGGTGTGGGAGAGAGAG	724				

1529 GTTGTAGCATGAGAGAGAGCCAGGACCCGGAAGAAGAGCCTGGGCCACGGCCACGT 1588
 725 GTTGTAGCATGAGAGAGAGCCAGGACCCGGAAGAAGAGCCTGGGCCACGGCCACGT 784
 1589 ATTAACCTCCAGGGGATATAGGTCCATCACTCTCTGGAGGCCCTCTTGGAGTGGAG 1648
 785 ATTAACCTCCAGGGGATATAGGTCCATCACTCTCTGGAGGCCCTCTTGGAGTGGAG 844
 1649 AGCACTCAGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1708
 845 AGCACTCAGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904
 1709 GAGCCTTCAGAGCCTTCCTCCACAGCTTCCAAAGAGACCAAGAGAGCCAGAGTGGACAG 1768
 905 GAGCCTTCAGAGCCTTCCTCCACAGCTTCCAAAGAGAGCCAGAGTGGACAG 964
 1769 GAGCCTTCAGAGCCTTCCTCCACAGCTTCCAAAGAGAGCCAGAGTGGACAG 1828
 965 GAGCCTTCAGAGCCTTCCTCCACAGCTTCCAAAGAGAGCCAGAGTGGACAG 1024
 1829 GTGGCCAAACCGAGACCCAGAGCCTTCAGAGAGCAAGAGAGAGGAGGAGGAGGAGGAG 1888
 1025 GTGGCCAAACCGAGACCCAGAGCCTTCAGAGAGCAAGAGAGAGGAGGAGGAGGAGGAG 1084
 1889 GAGCCTTCAGATTCCTCTACGCCCCAGGTTCCGAGAGGAGTGTAGACAGAGCCAGCGTGCAT 1948
 1085 GAGCCTTCAGATTCCTCTACGCCCCAGGTTCCGAGAGGAGTGTAGACAGAGCCAGCGTGCAT 1144
 1949 GACAGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2008
 1145 GACAGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
 2009 AAGAGATTCACAAACTCTTGGAGAAACACCTCAGCTGTGTTGCCGACACATTCCTCT 2068
 1205 AAGAGATTCACAAACTCTTGGAGAAACACCTCAGCTGTGTTGCCGACACATTCCTCT 1264
 2069 CAGCTCCGCCCATACCCGCTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAG 2128
 1265 CAGCTCCGCCCATACCCGCTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAG 1324
 2129 CAGCTCCGCCCATACCCGCTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAG 2188
 1325 CAGCTCCGCCCATACCCGCTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAG 1384
 2189 GCGTGTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2248
 1385 GCGTGTCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444
 2249 TGGGGGCAACAGACCCCTAGTTCATCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 2308
 1445 TGGGGGCAACAGACCCCTAGTTCATCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1504
 2309 CGGATGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2368
 1505 CGGATGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1564
 2369 CTGACCT 2428
 1565 CTGACCT 1624
 2429 CCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2488
 1625 CCCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1684
 2489 TCGGTTCCCTCATCATCT 2548
 1685 TCGGTTCCCTCATCATCT 1744
 2549 GACACACGCGCAGATCAACAGGAGCAGCAGGAGCAGAGATAGTCTTTTGTCTCTCTCTCT 2608
 1745 GACACACGCGCAGATCAACAGGAGCAGCAGGAGCAGAGATAGTCTTTTGTCTCTCTCTCT 1804
 2609 CTTCTGCTAGTCACTTTTTCATAGCTTACAGATAGTCTGGCTTGTCTCTCTCTCTCTCTCT 2668

1805 CCTGTGCTAGTCACTTTTTCATATGCTTACAGATGTGCTTGTCTGAGAGATTAATA 1864
 2669 CACATTTTCATTAATAAAAAAAAAAAAAAAAAAAAAA 2704
 1865 CACATTTTCATATTTTGTATTAATAAAAAAAAAAAAAA 1900
 RESULT 7
 BC004210 1905 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION
 Homo sapiens, similar to hypothetical protein FLJ10928, clone
 MGC:4436 IMAGE:2958967, mRNA, complete cds.
 ACCESSION
 BC004210
 VERSION
 BC004210.1 GI:13278902
 SOURCE
 HOMO SAPIENS
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 1905)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Ruben Laboratory
 DNA Sequencing by: Institute for Systems Biology
 contact: <http://www.systemsbio.org>
 contact: amadansystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 10 Row: 9 Column: 1.
 Location/Qualifiers
 1. 1905
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:4436 IMAGE:2958967"
 /tissue="type="Kidney, renal cell adenocarcinoma"
 /clone_lib="NIH-MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: pORF7"
 272..1171
 /codon_start=1
 /product="Similar to hypothetical protein FLJ10928"
 /protein_id="AAH04210.1"
 /db_xref="GI:13278903"
 /translation="MAREIDNFPYRFTYHNVRVMDSESAQLPHMKETRRFIEARA
 OSHVIVCKMGSRSAAVLAAMKQECSEOLRHOVEELRPLARNPGLQLOI
 YOGITLARSQSHVMEKQVGVSPREHAPAPSTPPRPPVPEPGEGERKVVGMESQ
 APKEPGRPRINLNGVMSRISLERSLESTSTSMPEVSESHSHEPILOPFP
 OLARKGOOVDRGPOPALKRSQSVVTLQGSAAVANKRQAFQEDQSGGCGGCEP
 STPRKVVVROASVHDSGESEA"
 BASE COUNT
 418 a 627 c 517 g 343 t
 ORIGIN
 Query Match 61.6%; Score 1665.2; DB 9; Length 1905;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 8; Indels 99; Gaps 2;
 907 GAGCAGCGATCCGTGTGAGCTGTGAAGCTGTGGAGTCACTGACCGGAGAGATGTC 966
 6 GAGAGGCGATCCGTGCT AGCTGTGAAGCTGTGGAGTCACTGACCGGAGAGATGTC 64

[illegible][illegible]

ORIGIN

Query Match 52.4%; Score 1416; DB 6; Length 1416;
 Best Local Similarity 100.0%; Pred No. 4.2e-283;
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 ATGGCCCTGGTACAGTACAGCCGTTGCCCGCCGAGCGCGCCCTCCACGCCCTGGGG 153
 Db 1 ATGGCCCTGGTACAGTACAGCCGTTGCCCGCCGAGCGCGCCCTCCACGCCCTGGGG 60
 OY 154 CCCTGGGACAGCGCGCTCCAGCGAGAGTGCAGTCCAGGAGGAGAGACTTTGGGTG 213
 Db 61 CCCTGGGACAGCGCGCTCCAGCGAGAGTGCAGTCCAGGAGGAGAGACTTTGGGTG 120
 OY 214 CTCCGTGGGCTGTCTGGGAGTGCAGAGTGCAGGAGGAGCAATGATGATGAGCAGAGGCC 273
 Db 121 CTCCGTGGGCTGTCTGGGAGTGCAGAGTGCAGGAGGAGCAATGATGATGAGCAGAGGCC 180
 OY 274 AGTTCTGAGCCAAACAGAGAGGCCCGAGTGCAGAGAGGCTCCAGCGGGACACAGAC 333
 Db 181 AGTTCTGAGCCAAACAGAGAGGCCCGAGTGCAGAGAGGCTCCAGCGGGACACAGAC 240
 OY 334 TTGGGGCAAGATCCCGAGTCCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 393
 Db 241 TTGGGGCAAGATCCCGAGTCCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 OY 394 ATGGTACAGCTGTGAGGCGCGAGAGTGCAGTCCGCTGGCAGCCAGCTGAGAGGAGCC 453
 Db 301 ATGGTACAGCTGTGAGGCGCGAGAGTGCAGTCCGCTGGCAGCCAGCTGAGAGGAGCC 360
 OY 454 CGGCTTCCCGGCTCCGCTACCTGCTGTGTAGTTCTACACAGAGAGAGAGAGTCTGAGC 513
 Db 361 CGGCTTCCCGGCTCCGCTACCTGCTGTGTAGTTCTACACAGAGAGAGAGAGTCTGAGC 420
 OY 514 CAGGATGAGAGGAGTCTGAGGCGGTGGATTTCCCTGACAGCAGCTCCGAGTGCAGC 573
 Db 421 CAGGATGAGAGGAGTCTGAGGCGGTGGATTTCCCTGACAGCAGCTCCGAGTGCAGC 480
 OY 574 CTGGGCTGTGTCTTGGCCCTCTGAGTGCAGACCCAGGTGTACTGATGAGAGCGGGGC 633
 Db 481 CTGGGCTGTGTCTTGGCCCTCTGAGTGCAGACCCAGGTGTACTGATGAGAGCGGGGC 540
 OY 634 TTTCAGCTGAGCTGTGGTGGGCAAGCCGGATCTTCAAGCCCATCTCATTCAGACCATG 693
 Db 541 TTTCAGCTGAGCTGTGGTGGGCAAGCCGGATCTTCAAGCCCATCTCATTCAGACCATG 600
 OY 694 TGGGCAACACTCCAGGTATTGCACCAAGCATGTAGGAGAGCTTAGGAGGAGGAGGAGGAGG 753
 Db 601 TGGGCAACACTCCAGGTATTGCACCAAGCATGTAGGAGAGCTTAGGAGGAGGAGGAGG 660
 OY 754 CCGGGTGGCAGTGCCTTACCTGGGCGCAGCCACTACCAAGAGAGTGAATCCGAACAG 813
 Db 661 CCGGGTGGCAGTGCCTTACCTGGGCGCAGCCACTACCAAGAGAGTGAATCCGAACAG 720
 OY 814 AGCTGCTCAATGATGAGAGGAGCTATGCGCAGCTGAGAGTCTGCGGCTCCAGCGCC 873
 Db 721 AGCTGCTCAATGATGAGAGGAGCTATGCGCAGCTGAGAGTCTGCGGCTCCAGCGCC 780
 OY 874 GAGCTGGGCGGCTCTCAGAAACAGAGAGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGG 933
 Db 781 GAGCTGGGCGGCTCTCAGAAACAGAGAGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGG 840
 OY 934 AAATGTTGATGTCAATGAGCTGAGAGTGTCACTTCCAAAGAGATCCGACAGCTCTG 993
 Db 841 AAATGTTGATGTCAATGAGCTGAGAGTGTCACTTCCAAAGAGATCCGACAGCTCTG 900
 OY 994 GAGCTGGGCGGCTCTCAGAAACAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1053
 Db 901 GAGCTGGGCGGCTCTCAGAAACAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 960
 OY 1054 CTGCTGGTGGAGCAGCGGAGACGAGCTCCCGGATCTTCCCGCAGCTCTACTGGCTCA 1113
 Db 961 CTGCTGGTGGAGCAGCGGAGACGAGCTCCCGGATCTTCCCGCAGCTCTACTGGCTCA 1020

OY 1114 GAGTGGAGCGCAACCACTGGAGAGTGCAGAGAGAGGAGTGCAGAGTCTTGAAC 1173
 Db 1021 GAGTGGAGCGCAACCACTGGAGAGTGCAGAGAGAGGAGTGCAGAGTCTTGAAC 1080
 OY 1174 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTGGCCTC 1233
 Db 1081 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTGGCCTC 1140
 OY 1234 TGGGATGAGAGTGGCGCCAGTGGCTGCGCAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 1293
 Db 1141 TGGGATGAGAGTGGCGCCAGTGGCTGCGCAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 1200
 OY 1294 GCTCAAGAGCAGAGGAGCCAGCGTGTGTCCATGCAAGATGGGCGTACAGCGCTCA 1353
 Db 1201 GCTCAAGAGCAGAGGAGCCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
 OY 1354 GCGGCCACAGTGTGGCTTATGCTGATGAGAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1413
 Db 1261 GCGGCCACAGTGTGGCTTATGCTGATGAGAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1320
 OY 1414 CGCCAGTGCAGAGTGGCGCCATGCGCCAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1473
 Db 1321 CGCCAGTGCAGAGTGGCGCCATGCGCCAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1380
 OY 1474 CAGATCTACAGGAGCATCTGACGGGCGAGACCTGA 1509
 Db 1381 CAGATCTACAGGAGCATCTGACGGGCGAGACCTGA 1416

RESULT 9
 AB072360 1416 bp mRNA linear PRI 29-JAN-2002
 LOCUS Homo sapiens mRNA for hSSH-3, complete cds.
 DEFINITION
 AB072360
 ACCESSION
 VERSION
 AB072360.1 GI:18376668
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens cDNA to mRNA.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 Niva, R., Negata-Ohashi, K., Takeichi, M., Mizuno, K. and Uemura, T.
 Control of actin reorganization by Slingshot, a family of
 phosphatases that dephosphorylate Arp/cofilin
 Cell 108 (2), 233-246 (2002)

JOURNAL
 MEDLINE
 21822082
 2 (bases 1 to 1416)

REFERENCE
 Niva, R., Negata-Ohashi, K., Hay, B. A., Takeichi, M., Mizuno, K. and
 Uemura, T.

TITLE
 Direct Submission
 JOURNAL
 Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular
 Genetics, The Institute for Virus Research, Kyoto University;
 Shogoin-Kawara-cho 53, Sakyo-ku, Kyoto 606-8507, Japan
 (E-mail: tuemura@virus.kyoto-u.ac.jp, Tel: 81-75-751-4031,
 Fax: 81-75-751-3989)

FEATURES
 Location/Qualifiers
 1..1416
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..1416
 /gene="hSSH-3"
 1..1416
 /gene="hSSH-3"
 /note="a putative dual specific phosphatase"

CDS
 /gene="hSSH-3"
 /product="hSSH-3"
 /note="a putative dual specific phosphatase"

/codon_start=1
 /product="hSSH-3"
 /protein_id="BAB84119.1"
 /db_xref="GI:18376669"
 /translation="MALVTVSRSPGSGASTPVGPMDQAVORRSRLOROSPAVLKGA
 VGLDGGDNDAVEASSPTPEKASPEELHDDQIDPFGSGSSPQKOEORHLHLMY
 QILRPDDIRILAOLEAPRPRLRYLTVSTREGSLSDDETVLLGVDPDSSSPSCT

LGLVPLMSDPTQVYLIDGGESVYSGOSRIFKFIOTIMATLQVLHQAECNALGSG
LVPGSALITWASHQERLNSQSLNEMTANLADLESIRPSPACGSSSEDOEMQAIR
AEIKVLDVDSLEBSEVTSKEIRLOMSKAGSSSTSVSSTTRCCMWSHSGBRATRP
HLVYGSMMANLEIOLRNRYTHLNMARENDNPPRFTYHNRLINDEBSAQLPPIH
KEHRTFEARAQSTHYLVHCKMGVSRSAATVLAAMKQIESLEQALRHVQELRPTA
RPNRGLRQLOIYOGILHART"

BASE COUNT 288 a 449 c 449 g 230 t

ORIGIN

Query Match 51.5%; Score 1391.4; DB 9; Length 1416;
Best Local Similarity 99.8%; Pred. No. 5.3e-278;
Matches 1414; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 94 ATGGCCCTGCTACAGTATGAGCGCTTCGCCCCGGGACAGCGGGCTTCACAGCCGCTGGGG 153
DB 1 ATGGCCCTGCTACAGTATGAGCGCTTCGCCCCGGGACAGCGGGCTTCACAGCCGCTGGGG 60

QY 154 CCTGGGACACAGGCGGCTCAGCGAAGAGTGCATCCAGCGAAGCAGAGCTTTCGGTG 213
DB 61 CCTGGGACACAGGCGGCTCAGCGAAGAGTGCATCCAGCGAAGCAGAGCTTTCGGTG 120

QY 214 CTCGCTGGGCGTGTCTCTGGACTGAGATGAGAGGGACATGATGATGACAGAGGCC 273
DB 121 CTCGCTGGGCGTGTCTCTGGACTGAGATGAGAGGGACATGATGATGACAGAGGCC 180

QY 274 AGTTCTGAGCCACAGAGAAGGCCCGAGTGCAGAGAGCTCCAGGGGACCAGACAGC 333
DB 181 AGTTCTGAGCCACAGAGAAGGCCCGAGTGCAGAGAGCTCCAGGGGACCAGACAGC 240

QY 334 TTCCGGGCAAGATCCACAGATGCCAGAAAGCAGAGAGCAGAGGACACCTGCACCTC 393
DB 241 TTCCGGGCAAGATCCACAGATGCCAGAAAGCAGAGAGCAGAGGACACCTGCACCTC 300

QY 394 ATGTGACAGCTGTGAGAGGCCCGAGATGACATCCGCTGGGAGCCAGCTGGAGGACACC 453
DB 301 ATGTGACAGCTGTGAGAGGCCCGAGATGACATCCGCTGGGAGCCAGCTGGAGGACACC 360

QY 454 CGGCTCCCGGCGCTCCGCTACCTGCTGATTTCTATACAGAGAGAGAGTCTGAGC 513
DB 361 CGGCTCCCGGCGCTCCGCTACCTGCTGATTTCTATACAGAGAGAGAGTCTGAGC 420

QY 514 CAGATGAGACGCTCTCTGGGCGTGAATTTCCCTGACAGACGCTCCCGACGTCACCC 573
DB 421 CAGATGAGACGCTCTCTGGGCGTGAATTTCCCTGACAGACGCTCCCGACGTCACCC 480

QY 574 CTGGGCGTGTCTTCCGCTCTGAGTACACCCGAGTGTCTTGTGATGAGAGACCGGGCC 633
DB 481 CTGGGCGTGTCTTCCGCTCTGAGTACACCCGAGTGTCTTGTGATGAGAGACCGGGCC 540

QY 634 TTACGCTGAGCTGTGAGGCGAAAGCGGATCTTCAAGCCCATCTCCATCCAGACATG 693
DB 541 TTACGCTGAGCTGTGAGGCGAAAGCGGATCTTCAAGCCCATCTCCATCCAGACATG 600

QY 694 TGGGCGACACTCCAGATATTCACCAAGCATGTGAGGAGCGCTCTAGGAGCGGCTTGT 753
DB 601 TGGGCGACACTCCAGATATTCACCAAGCATGTGAGGAGCGCTCTAGGAGCGGCTTGT 660

QY 754 CCGGCTGACAGTCCCTACCTGAGGCGACCACTACAGAGAGAGCTCAACTCCGAAGAG 813
DB 661 CCGGCTGACAGTCCCTACCTGAGGCGACCACTACAGAGAGAGCTCAACTCCGAAGAG 720

QY 814 AGTCCCTCAATGATGATGAGAGCGCTATGAGCGACCTGAGATCTTCCGCGCTCCAGGCC 873
DB 721 AGTCCCTCAATGATGATGAGAGCGCTATGAGCGACCTGAGATCTTCCGCGCTCCAGGCC 780

QY 874 GAGCTGGCGGCTCTCAAGACAGAGAGCAGATGAGAGCGGATCCGTGTGAGCTGTG 933
DB 781 GAGCTGGCGGCTCTCTCAAGACAGAGAGCAGATGAGAGCGGATCCGTGTGAGCTGTG 840

QY 934 AAAGTGTGATGATGATGAGAGCTGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTG 993
DB 841 AAAGTGTGATGATGATGAGAGCTGAGAGTGTCACTTCCAAAGAGATCCGAGGCTGTG 899

QY 994 GAGCTGCCCTGGGGCTCCCTCCAGCAGTACCTGATCTTCATCGACACAGATGCTG 1053
DB 900 GAGCTGCCCTGGGGCTCCCTCCAGCAGTACCTGATCTTCATCGACACAGATGCTG 959

QY 1054 CTGCTGTGGGCACAGC -GGAGCCAGGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTC 1112
DB 960 CTGCTGTGGGCACAGCAGGGGAGCCGAGGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTC 1019

QY 1113 AGACTGCAACCGACCAACCTGAGAGAGCTGACAGAGAAACAGAGGCTACCCACATCTTGA 1172
DB 1020 AGAGTGAACCGACCAACCTGAGAGAGCTGACAGAGAAACAGAGGCTACCCACATCTTGA 1079

QY 1173 CATGCCCCGGGAGATTGACAACTTACCTGAGGCGCTTACCTACCTCAATGTGCGCT 1232
DB 1080 CATGCCCCGGGAGATTGACAACTTACCTGAGGCGCTTACCTACCTCAATGTGCGCT 1139

QY 1233 GCGTCAAGAGACACAGGCGACCCAGCTGTGTCACCTGCAAGATGAGGCGCTCAGCGCTC 1352
DB 1200 GCGTCAAGAGACACAGGCGACCCAGCTGTGTCACCTGCAAGATGAGGCGCTCAGCGCTC 1259

QY 1353 AGCGGCACAGTCTGCGCTATGCTGATGAGAGAGTACGAATGACAGCTGAGACAGGCTC 1412
DB 1260 AGCGGCACAGTCTGCGCTATGCTGATGAGAGAGTACGAATGACAGCTGAGACAGGCTC 1319

QY 1413 GCGCCAGTGCAGAGCTCCGCGCCATGCGCCGCAACCTGCGCTTCTGCGCCAGCT 1472
DB 1320 GCGCCAGTGCAGAGCTCCGCGCCATGCGCCGCAACCTGCGCTTCTGCGCCAGCT 1379

QY 1473 GCAGATCTTACCAAGGCGCATCTGAGGCGCAGAACCTGA 1509
DB 1380 GCAGATCTTACCAAGGCGCATCTGAGGCGCAGAACCTGA 1416

RESULT 10
AX398768 1755 bp DNA linear PAT 27-MAY-2002
LOCUS AX398768
DEFINITION Sequence 4 from Patent W00220732.
ACCESSION AX398768
VERSION AX398768.1 GI:21261303
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Liou,J.R.
TITLE Regulation of human map kinase phosphatase-1-like enzyme
JOURNALS Patent: WO 0220732-A 4 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1. 1755
Location/Qualifiers
BASE COUNT 350 a 577 c 505 g 323 t
ORIGIN

Query Match 49.4%; Score 1337; DB 6; Length 1755;
Best Local Similarity 94.7%; Pred. No. 1e-266;
Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

QY 1174 ATGGCCCGGAGATTGACAACTTACCTGAGCGCTTACCTACCAATGTGCGCTC 1233
DB 322 ATGGCCCGGAGATTGACAACTTACCTGAGCGCTTACCTACCAATGTGCGCTC 381

QY 1234 TGGGATGAGAGTGGCGCCAGTGTGCGGACATGGAAGAGAGAGCAGCGCTCAATTGAG 1293
DB 382 TGGGATGAGAGTGGCGCCAGTGTGCGGACATGGAAGAGAGAGCAGCGCTCAATTGAG 441

QY 1294 GCTGCAAGAGCACAGGCGCACCCAGCTGCTGCTCACTGCAAGATGGCGGTGCA 1353

```

|||||
Db 442 GCTGCAAGAGACAGAGGACACCACCTGCTGTCACATGCAAGATGGGCTCAGCCGCTCA 501
OY 1354 GGGGCAAGATGCTGGCCATGATGCAAGAGATGAGATGAGCTGGAGAGGCGCTG 1413
Db 502 GGGGCAAGATGCTGGCCATGATGCAAGAGATGAGATGAGCTGGAGAGGCGCTG 561
OY 1414 GGGCAGATGACAGAGCTCGGGCCATCGCCCGCCCAACCCCTGGCTTCTTGGCCAGCTG 1473
Db 562 GGGCAGATGACAGAGCTCGGGCCATCGCCCGCCCAACCCCTGGCTTCTTGGCCAGCTG 621
OY 1474 CAGATCTTACAGAGGATCTCTGACGGCCAGAACCTTGAAGGTTGGTGGGAGAGAGTTCT 1533
Db 622 CAGATCTTACAGAGGATCTCTGACGGCCAGAACCTTGAAGGTTGGTGGGAGAGAGTTCT 681
OY 1534 AGGCTGGAAGAGAGCCAGAGGACCCCGAAGAGAGAGCTTGGGCGACGCGCATATAA 1593
Db 682 AGGCTGGAAGAGAGCCAGAGGACCCCGAAGAGAGAGCTTGGGCGACGCGCATATAA 741
OY 1594 CCTCGAGAGGATCATGAGTTCATGAGTCTTGGAGCCCTCTTGGAGCTGGAGAGCAC 1653
Db 742 CCTCGAGAGGATCATGAGTTCATGAGTCTTGGAGCCCTCTTGGAGCTGGAGAGCAC 801
OY 1654 CTCAGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1713
Db 802 CTCAGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
OY 1714 TCTGAGAGCCCTTCCAGACCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1773
Db 862 TCTGAGAGCCCTTCCAGACCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
OY 1774 TCAGGCTGCTGAAATGCTGAGTCCGAGTCACTGCTTCCAGAGGAGTCCCTGCTG 1833
Db 922 TCAGGCTGCTGAAATGCTGAGTCCGAGTCACTGCTTCCAGAGGAGTCCCTGCTG 981
OY 1834 CAACCGGAGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1893
Db 982 CAACCGGAGAGGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041
OY 1894 CTGATTTTCTTACAGCCAGGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1953
Db 1042 CTGATTTTCTTACAGCCAGGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101
OY 1954 TGGAGAGAGAGGAGGCTGAGCTTCAACATGCTTCCAGAGGCTTCCCTTCAAGAG 2013
Db 1102 TGGAGAGAGAGGAGGCTTCAACATGCTTCCAGAGGCTTCCCTTCAAGAG 1161
OY 2014 GATCCACATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2073
Db 1162 GATCCACATCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1221
OY 2074 CGGCCCCATACCCGCTTACAGAGCTTCCAGAGGCTTCCAGAGGCTTCCAGAGGCT 2133
Db 1222 CGGCCCCATACCCGCTTACAGAGCTTCCAGAGGCTTCCAGAGGCTTCCAGAGGCT 1281
OY 2134 CACCCCTGTACTACAGCTTCACTCTTCAAGCTTAACTTAACTTAACTTAACTTAA 2193
Db 1282 CACCCCTGTACTACAGCTTCACTCTTCAAGCTTAACTTAACTTAACTTAACTTAA 1341
OY 2194 TCCAGAGGCTTCAAGATTTTCTTAACTTGAATGCTTGAAGAGAGAGAGAGAGAG 2253
Db 1342 TCCAGAGGCTTCAAGATTTTCTTAACTTGAATGCTTGAAGAGAGAGAGAGAGAG 1401
OY 2254 GCAACAGAGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 2313
Db 1402 GCAACAGAGAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1434
OY 2314 TGAAGAGAGAGCTTCCGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2373
Db 1435 -----CCTTCA 1441
OY 2374 CCTCTGCTCTCTCCAGATTCATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2433
|||||

```

```

Db 1442 CCTCTGTCTCTCCAGTTTCTTCTTGAAGACACAGAGAGAGAGAGAGAGAGAGAG 1501
OY 2434 AAAGCAGAGAGAGATCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2493
Db 1502 AAAGCAGAGAGAGATCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561
OY 2494 TCCCTTATTCACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2553
Db 1562 TCCCTTATTCACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
OY 2554 CACGCGAGATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2613
Db 1622 CACGCGAGATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
OY 2614 GCGTAGTCAAGTTTTCATACCTTACAGTATGCTGCTTGTGTAAGAGAGAGAGAGAT 2673
Db 1682 GCGTAGTCAAGTTTTCATACCTTACAGTATGCTGCTGCTGTAAGAGAGAGAGAGAT 1741
OY 2674 TTTTCATA 2680
Db 1742 TTTTCATA 1748

RESULT 11
AK001790 1755 bp mRNA linear PRI 01-AUG-2002
LOCUS AK001790
DEFINITION Homo sapiens cDNA FLJ10928 fis, clone OVARC100473, weakly similar
to DUAL SPECIFICITY PROTEIN PHOSPHATASE 3 (EC 3.1.3.48) (EC
3.1.3.16).
ACCESSION AK001790
VERSION AK001790.1 GI:7023282
KEYWORDS oligo cloning; fls (full insert sequence).
SOURCE Homo sapiens ovarian cancer cDNA to mRNA, clone_1lb:OVARC1
clone:OVARC100473.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Ishibashi,T., Fujimori,K.,
Tanai,H., Kimata,M., Matanabe,M., Hiraoka,S., Ishi,S., Kawai,Y.,
Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahori,K.,
Masuko,Y. and Kanehori,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1755)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source 1..1755
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC100473"
/issue_type="ovarian cancer"
/clone_1lb="OVARC1"
/note="cloning vector: pME18FLJ3"
232..657
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91913.1"
/db_xref="GI:7023283"
/translation="MEGTMMDQORPVLSDQHPFILNSSPAHSPAREIDNFPYPERFT
YHNVRLMDEBSQQLPHWKETHRFLEARAGCTHYLVNCKMGVSNATVLAAMKQY

```

BASE COUNT 350 a 577 c 505 g 323 t
 ORIGIN

Query Match 49.4%; Score 1337; DB 9; Length 1755;
 Best Local Similarity 94.7%; Pred. No. 1e-266;
 Matches 1427; Conservative 0; Mismatches 0; Indels 80; Gaps 1;

1174 ATGGCCGGGAGATTGACAACTTCTACCCCTGAGCGCTTACCTACCAACAATGGCGCTC 1233
 322 ATGGCCGGGAGATTGACAACTTCTACCCCTGAGCGCTTACCTACCAACAATGGCGCTC 381
 1234 TGGATGAGAGAGTCCGCCAGCTGCTGCGGCACTGGAAGAGAGCACCCTTCAATTGAG 1293
 382 TGGATGAGAGAGTCCGCCAGCTGCTGCGGCACTGGAAGAGAGCACCCTTCAATTGAG 441
 1294 GCTGGAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1353
 442 GCTGGAAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
 1354 GCGCCAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
 502 GCGCCAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
 1414 GCGCCAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
 562 GCGCCAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
 1474 CAGATCTACAGGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1533
 622 CAGATCTACAGGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
 1534 AGGCAATGAGAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1593
 682 AGGCAATGAGAGAGCAGAGGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741
 1594 CCTCCAGAGGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1653
 742 CCTCCAGAGGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
 1654 CTGACAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1713
 802 CTGACAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 861
 1714 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1773
 862 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 921
 1774 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1833
 922 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
 1834 CAACCGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1893
 982 CAACCGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1041
 1894 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1953
 1042 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
 1954 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2013
 1102 TCTGAGAGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1161
 2014 GATTCACAACTCTGTTGGAAGAACCTTCACAGTCTGTTGCGGCACACATCTCTCAGCT 2073
 1162 GATTCACAACTCTGTTGGAAGAACCTTCACAGTCTGTTGCGGCACACATCTCTCAGCT 1221
 2074 CGGCCCCAATACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2133
 1222 CGGCCCCAATACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
 2134 CACCCCTGTCTACTACAGCTCTCTCTACAGCTTAACTGATCCAGGCGCCATGTCTGCTG 2193

1282 CACCCCTGTCTACTACAGCT 1341
 2194 TCCAGAGGCTCAAGACTTCT 2253
 1342 TCCAGAGGCTCAAGACTTCT 1401
 2254 GCAACAGCACCCTAGTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2313
 1402 GCAACAGCACCCTAGTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1434
 2314 TGAAGAGAGCTTCCCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2373
 1435 -----CCCTGCA 1441
 2374 CCTCTGCT 2433
 1442 CCTCTGCT 1501
 2434 AAAGCAGGAGAGAGTCT 2493
 1502 AAAGCAGGAGAGAGTCT 1561
 2494 TCCCTCATCTACCT 2553
 1562 TCCCTCATCTACCT 1621
 2554 CAGGCGAGTCAAGAGGAG 2613
 1622 CAGGCGAGTCAAGAGGAG 1681
 2614 GCGTATGATCTTTCATAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2673
 1682 GCGTATGATCTTTCATAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1741
 2674 TTTCTATA 2680
 1742 TTTCTATA 1748

RESULT 12
 AK000522
 LOCUS 2064 bp mRNA linear PRI 22-FEB-2000
 DEFINITION Homo sapiens cDNA FLJ20515 fls, clone KAT09889.
 ACCESSION AK000522.1 GI:7020673
 VERSION AK000522.1
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III CDNA to mRNA, clone_11b:KAT clone:KAT09889.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2064)
 TITLE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
 JOURNAL Direct Submission
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@iims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
source
Location/Qualifiers
1. 2064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT09889"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_lib="KAT"
/note="cloning vector pME18SFL3"
147. 1331
/note="unnamed protein product"
/protein_id="BAA9128.1"
/db_xref="gi:7020674"
/translation="MALVTSRSPGSGASFPVGPMDQAVORSHRLOROSFAVLGCA
VLGDDGDNDNAAPASSEPEKAPSEELHGDPTDPOGSGSPKOEOROHILMV
QLRPQDIDRLAOLEAPRPRLRLYLIVSTREGSLSDDEYVLGVFPDSSPSCT
LGLVPLMSDQVYLDGDSFVTSGGSRLEKPLISITQTMATLQVLAHQAALGSG
LVPGSALTMAVSHYOERLNSQSCINEMTAMADLESLPPSAPEGSSBOEOMEQAIK
AELMKVLELESTSESDMEVFSHSHSHEEPLOFPOLARTKGQOYDRGPOLAKS
ROSVTLGSAVAVNRTOAFQEQGGGQGEPCISSPFRKVVROASVHDSGEEG
EA"

CDS
BASE COUNT 421 a 671 c 605 g 367 t
ORIGIN
Query Match 47.5%: Score 1285.6; DB 9: Length 2064;
Best Local Similarity 73.9%: Pred. No. 4.5e-256;
Matches 1998; Conservative 0; Mismatches 9; Indels 697; Gaps 1:

QY 1 CGGCTTCTGTCGTCGCGGGGTCAGAGCTGTCCGCGGGTTAGGAGAGGCGCTGCC 60
|||
DB 54 CGTCTTCTGTCGTCGCGGGGTCAGAGCTGTCCGCGGGTTAGGAGAGGCGCTGCC 113
QY 61 CGGTGCGAGCCAGAGTGTGTCGCGGCTGCTCATGGCCCTGTGTCACAGAGAGCGTTGC 120
|||
DB 114 CGGTGCGAGCCAGAGTGTGTCGCGGCTGCTCATGGCCCTGTGTCACAGAGAGCGTTGC 173
QY 121 CCCCCGCGAGCGCGGCTTCACGCCCCGTGGGGCCCTGGAGCAGAGCGGTCCAGCGAAG 180
|||
DB 174 CCCCCGCGAGCGCGGCTTCACGCCCCGTGGGGCCCTGGAGCAGAGCGGTCCAGCGAAG 233
QY 181 AGTCGACTCAGAGGAGGAGGAGCTTTGGCGTGTCTCGTGGGGCTGTCTCGGAGCTGAG 240
|||
DB 234 AGTCGACTCAGAGGAGGAGGAGCTTTGGCGTGTCTCGTGGGGCTGTCTCGGAGCTGAG 293
QY 241 GATGAGAGGAGCATATGATGATGAGAGAGGAGCTTCTAGGCCAAGAGAGAGCGCCG 300
|||
DB 294 GATGAGAGGAGCATATGATGATGAGAGAGGAGCTTCTAGGCCAAGAGAGAGCGCCG 353
QY 301 AGTGAGAGGAGCTCCAGGGGGACACAGAGACTTCGGGCAAGGATCCAGAGTCCCCAG 360
|||
DB 354 AGTGAGAGGAGCTCCAGGGGGACACAGAGACTTCGGGCAAGGATCCAGAGTCCCCAG 413
QY 361 AAGCAGAGAGAGAGAGGAGGAGGAGCTGACCTGATGATGATGAGGAGGAGGAGGAGGAT 420
|||
DB 414 AAGCAGAGAGAGAGAGGAGGAGGAGCTGACCTGATGATGATGAGGAGGAGGAGGAGGAT 473
QY 421 GACATCGGCTGGAGCGGAGCTGAGGAGCACCAGGCTCCCGGCTCCCTACTGCTGCTG 480
|||
DB 474 GACATCGGCTGGAGCGGAGCTGAGGAGCACCAGGCTCCCGGCTCCCTACTGCTGCTG 533
QY 481 GTAGTTTCTACAGAGAGAGAGAGGTCGTGAGCCAGATGAGAGGTCCTCTCGGCGCTG 540
|||
DB 534 GTAGTTTCTACAGAGAGAGAGAGGTCGTGAGCCAGATGAGAGGTCCTCTCGGCGCTG 593
QY 541 GATTTCCCTGACAGAGCTCCCGAGCTGACACCTGGGCTGTGCTGCTGCTGCTGAGT 600
|||
DB 594 GATTTCCCTGACAGAGCTCCCGAGCTGACACCTGGGCTGTGCTGCTGCTGCTGAGT 653
QY 601 GACACCCAGAGTACTTATGATGAGAGGAGGAGGCTTACAGCGTGAAGTGGGCAAGC 660
|||
DB 654 GACACCCAGAGTACTTATGATGAGAGGAGGAGGCTTACAGCGTGAAGTGGGCAAGC 713

QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGCACACCCAGATATGACCAA 720
|||
DB 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGCGCACACCCAGATATGACCAA 773
QY 721 GCATGTGAGGAGCTCTGAGGAGCGGCTTGTACCGGCTGGCAGTGCCTCAGCTGGGCC 780
|||
DB 774 GCATGTGAGGAGCTCTGAGGAGCGGCTTGTACCGGCTGGCAGTGCCTCAGCTGGGCC 833
QY 781 AGCCACTACAGAGAGAGCTGAACCTCGACAGAGAGCTGCTCAATGATGAGAGGCTATG 840
|||
DB 834 AGCCACTACAGAGAGAGCTGAACCTCGACAGAGAGCTGCTCAATGATGAGAGGCTATG 893
QY 841 GCCGACCTGAGTCTCTGCGGCTCCCGACGCGGAGCCTGGCGGGTCTCTCAAGAACAGAG 900
|||
DB 894 GCCGACCTGAGTCTCTGCGGCTCCCGACGCGGAGCCTGGCGGGTCTCTCAAGAACAGAG 953
QY 901 CGATGAGAGCAGCGGATCCGCTGCTGAGCTGTGAAGTGTGATGATGATGAGTCTGAG 960
|||
DB 954 CGATGAGAGCAGCGGATCCGCTGCTGAGCTGTGAAGTGTGATGATGATGAGTCTGAG 992
QY 961 AGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGTGGGCTGCCCTCCCTCAG 1020
|||
DB 993 ----- 992
QY 1021 CAGTACCTGACTTCATGACAAACAGATGCTGTGCTGTGGCAGAGGAGCCGAGCC 1080
|||
DB 993 ----- 992
QY 1081 TCCCGATCTTCCCGACCTTACTGAGGCTCAGATGAGTGAACCAACCTGGAGAG 1140
|||
DB 993 ----- 992
QY 1141 CTGACAGAGAGAGAGGTCACCCACATCTTGAACATGCCCCGAGATTGACAACTTCTAC 1200
|||
DB 993 ----- 992
QY 1201 CCTGAGCGCTTCACTACCAATGTGGGCTGTGGGATGAGGAGTGGGCCAGCTGCTG 1260
|||
DB 993 ----- 992
QY 1261 CCGACTGGAAGAGAGCAGCAGCGCTTCATTGAGGCTGCAAGAGCAGAGGACCCAGCTG 1320
|||
DB 993 ----- 992
QY 1321 CTGCTCCTGCAAGATGGGCTTCAAGCCCTCAAGCCGACAGCTGTGCTATGCGATG 1380
|||
DB 993 ----- 992
QY 1381 AAGCAGTACGAATGACGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
|||
DB 993 ----- 992
QY 1441 GCCGCCCAACCTGCTGCTTCTGCGCAGCTGAGATGTACAGAGGATCTCTGAGGCGC 1500
|||
DB 993 ----- 992
QY 1501 AGAAGCTGAGGCTGTGGGAGAGAGAGGTTGAGCATGGAAGAGAGCAGGAGGAGCC 1560
|||
DB 993 ----- 992
QY 1561 GAAAGAGAGCCTGGGCGACGCGCAGCTATAAACCTCCGAGGAGTATGAGTCCATCAG 1620
|||
DB 993 ----- 992
QY 1621 TCTTGTGAGCCCTCTGTTGAGAGCTGAGAGACCTGAGAGACAGTGAATGCGAGAGT 1680
|||
DB 993 ----- 992
QY 1681 CTCTCTTCCACAGAGTCTTACATGAGAGGCTCTGAGGAGGCTTCTCCAGAGTTCAGAG 1740
|||
DB 1037 CTCTCTTCCACAGAGTCTTACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
|||
QY 1741 GACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800

|||||
Db 594 GATTCCCTGCACAGCTCCGCCAGCTGCACCCCTGGGCTTGTCCCTCTGGAGT 653
QY 601 GACACCCAGGTACTTAGATGAGAGGGGGCTTCAGCTGAGCTGTGGGCAAGC 660
Db 654 GACACCCAGGTACTTAGATGAGAGGGGGCTTCAGCTGAGCTGTGGGCAAGC 713
QY 661 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGGCAGACTCCAGGTATTGCACAA 720
Db 714 CGGATCTTCAAGCCCATCTCCATCCAGACATGTGGGGCAGACTCCAGGTATTGCACAA 773
QY 721 GCATGTGAGGAGCTCTTAGGAGCGGGCTTGTACCGGGTGGCAGTCCCTACCTGGGCC 780
Db 774 GCATGTGAGGAGCTCTTAGGAGCGGGCTTGTACCGGGTGGCAGTCCCTACCTGGGCC 833
QY 781 AGCCACTACGAGAGAGCTGAACCTCCAGACAGAGCTGCTCAATGATGAGGAGCTATG 840
Db 834 AGCCACTACGAGAGAGCTGAACCTCCAGACAGAGCTGCTCAATGATGAGGAGCTATG 893
QY 841 GCCGACCTGAGGTCTCTGCGGCTTCCAGCGCCGAGCCTGGCGGGTCTCAGAACAGAG 900
Db 894 GCCGACCTGAGGTCTCTGCGGCTTCCAGCGCCGAGCCTGGCGGGTCTCAGAACAGAG 953
QY 901 CAGATGAGAGCGGAGTCCGTGCTGAGCTGTGGAAGTGTGGATGTGATGACCTGGAG 960
Db 954 CAGATGAGAGCGGAGTCCGTGCTGAGCTGTGGAAGTGTGGATGTGATGACCTGGAG 992
QY 961 AGTGCATCTTCAAGAGAGATCCGCGAGGCTGTGAGTGTGGGCTGGGCTCCCTCCAG 1020
Db 993 ----- 992
QY 1021 CAGTACCTGACTTCATCGACACACAGATGCTGTGTGTGGCAGAGCGGAGCCAGGCC 1080
Db 993 ----- 992
QY 1081 TCCCGCATCTTCCCCACCTCTACTGCGGCTCAGAGTGGAAAGCAGCAAACTGGAGAG 1140
Db 993 ----- 992
QY 1141 CTGCAGAGAACAGGGTCAACCCACATCTTGAACATGGCCGGGAGATTGACAACTTCTAC 1200
Db 993 ----- 992
QY 1201 CCTGAGCCCTTACCTACCAACATGTGCGCTGTGGGATGAGAGATCGGCCAGCTGCTG 1260
Db 993 ----- 992
QY 1261 CCGCAGCTGGAAGAGAGCAGCCGTTTATGTAGGCTGCAAGAGCAGAGGGCAGCCAGGTG 1320
Db 993 ----- 992
QY 1321 CTGTCCACTGCAGAGATGGGGCTCAGCGCTCAGCGGCCACAGTGTGGCTATGCAATG 1380
Db 993 ----- 992
QY 1381 AAGCAGTACGAATGCACTTGAGCAGGCGCTGCGCCAGCTGACAGAGCTCCGCCCATC 1440
Db 993 ----- 992
QY 1441 GCCGCGCCCAACCTTGCTTCTGCGCAGCTGAGATCTACAGAGGATCTCTGAGGGCC 1500
Db 993 ----- 992
QY 1501 AGAACCTGAGGGTGTGGGAGAGAGAGTTGTAGCATGGAAGAGAGCCAGGACGCC 1560
Db 993 ----- 992
QY 1561 GAAGAAGAGCCTGGGCAAGCGCACGTATAAACCCTCGAGGGGTGATGAGGTCCATCAG 1620
Db 993 ----- 992
QY 1621 TCTTCTGAGGCCCTCTTGAGAGCTGGAGAGCACTCAGAGACCATGACATGCCAGAGT 1680
|||||

Db 993 -----TTGAGCTGGAGAGCACTCAGAGACCACTGACATGCCAGAGT 1036
QY 1681 CTTCCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACAGCTTGCAG 1740
Db 1037 CTTCCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACAGCTTGCAG 1096
QY 1741 GACCAAGGGAGCCAGCAGGTGAGAGAGGGGCTTCAGCTCCCTGGAAGTCCGCCAGTC 1800
Db 1097 GACCAAGGGAGCCAGCAGGTGAGAGAGGGGCTTCAGCTCCCTGGAAGTCCGCCAGTC 1156
QY 1801 AGTGTACCTCCAGAGGCACTGCGGTGGCCCAACCGGACCCAGGCTTCCAGAGCA 1860
Db 1157 AGTGTACCTCCAGAGGCACTGCGGTGGCCCAACCGGACCCAGGCTTCCAGAGCA 1216
QY 1861 GGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGCACTTCTCTACGCCAGGTTCCG 1920
Db 1217 GGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGCACTTCTCTACGCCAGGTTCCG 1276
QY 1921 GAAGGTGTGAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGGGCGAGGCTTGACCT 1980
Db 1277 GAAGGTGTGAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGGGCGAGGCTTGACCT 1336
QY 1981 CACACATGCCCCAGGCTCCCTGACACCTGAGAGAGATCCACAACCTCTTGAGAAACACCC 2040
Db 1337 CACACATGCCCCAGGCTCCCTGACACCTGAGAGAGATCCACAACCTCTTGAGAAACACCC 1396
QY 2041 TCAGCTGTGTGCGGACACACATTCCTCTCAGCTCGGCCCATACCCGTACAGGCTC 2100
Db 1397 TCAGCTGTGTGCGGACACACATTCCTCTCAGCTCGGCCCATACCCGTACAGGCTC 1456
QY 2101 ACCTCCACCCCTGTGCTACTAGGGCTTCACCTCCACCCCTTCACTACAGCTCACCTCC 2160
Db 1457 ACCTCCACCCCTGTGCTACTAGGGCTTCACCTCCACCCCTTCACTAGGCTCACCTCC 1516
QY 2161 TACAGCTTAAAGTCCACAGGCCATGTGCTGTCCAGAGGCTCAAGACTTCTTAAGTG 2220
Db 1517 TACAGCTTAAAGTCCACAGGCCATGTGCTGTCCAGAGGCTCAAGACTTCTTAAGTG 1576
QY 2221 GATGTGTAGAGGAGCTGAAGTACCTTGGGGGACACAGCAGCCCTAGTTCAATTCAC 2280
Db 1577 GATGTGTAGAGGAGCTGAAGTACCTTGGGGGACACAGCAGCCCTAGTTCAATTCAC 1636
QY 2281 CTCTAGCCCTGCACACTCACTGTGGCAGGAAATGAACAGAGCTTCCGCGCAAAAG 2340
Db 1637 CTCTAGCCCTGCACACTCACTGTGGCAGGAAATGAACAGAGCTTCCGCGCAAAAG 1696
QY 2341 GGTACGCTCCACACCCCGCCCTCCCTGACACTCTGTCTCTTCCAGTTCAATTCCT 2400
Db 1697 GGTACGCTCCACACCCCGCCCTCCCTGACACTCTGTCTCTTCCAGTTCAATTCCT 1756
QY 2401 GGACACAGCCAGGCGCAGCAACAGATGAGCCGCCCAAGAGGAGATCTCAGGCCCA 2460
Db 1757 GGACACAGCCAGGCGCAGCAACAGATGAGCCGCCCAAGAGGAGATCTCAGGCCCA 1816
QY 2461 GCCGCGGAGAGGTGGAAGGCTGGAGATCGCTTCCCTATCCACATCCAGCGGTCCAGG 2520
Db 1817 GCCGCGGAGAGGTGGAAGGCTGGAGATCGCTTCCCTATCCACATCCAGCGGTCCAGG 1876
QY 2521 TCTTGTGCTGTGCCAGAGCTCTGTGACACACAGCCAGATCAGAGGAGCAGGCGCA 2580
Db 1877 TCTTGTGCTGTGCCAGAGCTCTGTGACACACAGCCAGATCAGAGGAGCAGGCGCA 1936
QY 2581 GAGATAGTCTTCTTCTTCTTCTTCTTCTGAGCTGTGATGATGTTTTCATAGCTTACA 2640
Db 1937 GAGATAGTCTTCTTCTTCTTCTTCTTCTGAGCTGTGATGATGTTTTCATAGCTTACA 1996
QY 2641 GTATCTGGCTTTGTACTAGAAATATAACACATTTTCAATATTTTGGTTATTTTCAAAAAA 2700
Db 1997 GTATCTGGCTTTGTACTAGAAATATAACACATTTTCAATATTTTGGTTATTTTCAAAAAA 2056
QY 2701 AAAA 2704
Db 2057 AAAA 2060

RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	REMARK	COMMENT
BC028922	2736 bp	mRNA	linear	ROD 07-AUG-2002								
Mus musculus, similar to slingshot 3, clone MGC:25738												
IMAGE:3987714, mRNA, complete cds.												
BC028922												
BC028922.1	GI:20810382											
MGC.												
house mouse.												
Mus musculus												
Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.												
1 (bases 1 to 2736)												
Strausberg, R.												
Direct Submission												
Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA												
NIH-MGC Project URL: http://mgc.nci.nih.gov												
Contact: MGC help desk												
Email: cgabbs-rt@mail.nih.gov												
Tissue Procurement: Gilbert Smith, Ph.D.												
cDNA Library Preparation: Life Technologies, Inc.												
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)												
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center												
Center code: BCM-HGSC												
Web site: http://www.hgsc.bcm.tmc.edu/cdna/												
Contact: amg@bcm.tmc.edu												
Gunatentre, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.												
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov												
Series: IRAC Plate: 30 Row: 9 Column: 9												
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein.												
Location/Qualifiers												
1. 2736												
/organism="Mus musculus"												
/db_xref="taxon:10090"												
/map="CZECH II"												
/clone="MGC:25738 IMAGE:3987714"												
/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MTVV."												
/clone_1bp="NCI_CGAP_Lu29"												
/lab_host="DH10B"												
/note="Vector: pCMV-SPORT6"												
93. 2042												
/codon_start=1												
/product="Similar to slingshot 3"												
/protein_id="AAH28922.1"												
/db_xref="GI:20810383"												
/translation="MALVTSPSPASGSHSTPVGPTODRYVRRGRQLQRRSFAYLNGAVLGGDGSNVASDEAPMEPSGEDEPTDEOTDGGGLOSPKQVOKRHLMLWELIRAPRODILADLEAFRRPRLRYLLVLSGTGLSFEALLDQPPDSSHSTGLVPLPMSITQYLLDGDGFSVTSGGQSHIFRPVSIQTWMTLQVLRHQAECVALSGIVPGGSLAWATYVYOEKLNSDGCLNEMWMSDIESRPNAEBGASDEKMEQALIAELMOWYLSNDLDSVTSKEIRQALELRGL												

[illegible]

QY 1019 AGCAGTACCGTACTTTCATCGAACACAGATGCTGCTGCTGGCAGACGGGGACCGAG 1078
 Db 1009 AGCAGTACCGTACTTTCATCGAACACAGATGCTGCTGCTGGCAGACGGGGACCGAG 1068
 QY 1079 CCTCCCGATCTTCCCGACCTCTACCTGGGCTGAGTGGAGGACGACCAACCTGGAGG 1138
 Db 1069 CCTCCCGATCTTCCCGACCTCTACCTGGGCTGAGTGGAGTGGAGTGGAGTGGAGG 1128
 QY 1139 AGCTGAGAGGAAAGGGTTCACCCACATCTTGAACATGGCCGGGAGATTGACATCTCT 1198
 Db 1129 AACTTCAGAAAAACAGAGTAAATGATCATCTGACATGGCCGGGAGATTGACATCTCT 1188
 QY 1199 ACCGTGAGCGCTTTCACCAACAAATGTCGCTCTGGAGTGAAGATGCGCCAGCTGC 1258
 Db 1189 TCCCTGAGCGCTTTCACCAATTCATGCTGGAGTGAAGATGCGCCAGCTTTC 1248
 QY 1259 TGCCCATCTGGAAGAGACGACCGCTTTCATGAGCTGCAAGACACAGGCGCACCGAG 1318
 Db 1249 TGCCCATCTGGAAGAGACGACATTCATGAGAGATGCGACAGACGCGCGACTCGGG 1308
 QY 1319 TGCTGCTCCACTGCAAGATGGGCTGACCGCTGACGCGCGCAGTGTGGCTATGCCA 1378
 Db 1309 TGCTGCTCCACTGCAAGATGGGCTGACCGCTGACGCGCGCAGTGTGGCTATGCCA 1368
 QY 1379 TGAAGCATGACAAATGACAGCTGAGAGAGGCGCTGCGCCAGCTGCGGCCCA 1438
 Db 1369 TGAAGCATGACAAATGACAGCTGAGAGAGGCGCTGCGCCAGCTGCGGCCCA 1428
 QY 1439 TGCGCCGCGCCCAACCTGAGCTTCTGCGCGCAGCTGAGATACCGAGGCTATGACGG 1498
 Db 1429 TTGTAGCGCCCAACCGAGGCTTCTGCGCGCAGCTGAGATACCGAGGCTATGACGG 1488
 QY 1499 ----- 1498
 Db 1489 CCAGCGGAGAGACCACTGCTGGAGACAGAAAGTGGGTGCTCTCCAGAGAGACCCC 1548
 QY 1499 -----CCAGAACCTGAGGGGTGGGG 1520
 Db 1549 TGCGCGCTGAAGTTTCTACACCATTTGCACTTTCACCAAGGAGGAGGAGTGGGG 1608
 QY 1521 AGGAGAAAGTTTGAAGCATGAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1580
 Db 1609 AGGTGATGATGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1668
 QY 1581 GGGCAGCATAAACCTCGAGGGGCTATGAGTGCATGATCTTCTGAGAGGCTCTTGG 1640
 Db 1669 GGGCAGCATAAACCTCGAGGGGCTATGAGTGCATGATCTTCTGAGAGGCTCTTGG 1725
 QY 1641 AGCTGAGAGACACTGAGAGACAGTGCATGCGAGAGGCTCTCTCCAGAGAGCTT 1700
 Db 1726 AGTCAAGAGATCCCGAGAGGCGGAGGCTTGCAGAAAGTTTTC-----TCTT 1773
 QY 1701 CACATGAAGAGCTTTCAGAGGCTTCCAGAGCTTCAAGAGGAGGAGGAGGAGGAGGAG 1760
 Db 1774 CAGATGAAGAGCTTTCATTCATCTCTCTCAAGGCTTCAAGAGGAGGAGGAGGAGGAG 1833
 QY 1761 TGAGAGGGGGCTCGAGGCTGCGCTGAAGTCCGCGCATGATGTTTACCCTCGAGGCA 1820
 Db 1834 TCCGCAAGGGCTTGGCGCTGCGCTGAAGTCTGAGAGTCTGTTGCTTCAATAGTG 1893
 QY 1821 GTGCCGTGTGGGCAACCGAGGAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGG 1880
 Db 1894 GTGCCGTGTGGGCAACCGAGGAGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGG 1947
 QY 1881 GGCAGGAGAGGCGCTGATTTCTTACGCGCAGGTTCCGAGAGGAGGAGGAGGAGGAGG 1940
 Db 1948 AGGAGAGTGAAGCTGTGTATCTTCTTACCCCTTACGCTCGGAGAGGATGATGAGGAGG 2007
 QY 1941 GCGTGCATGACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1999
 Db 2008 GTGTAGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2067
 QY 2000 CTGACACTGAAGAGATCCAAACTCTTGGG-GAAAACACCTTACGCTGTGTTGGCCGAC 2058

Db 2068 AACTTAATGCGAGCTGTGTCAGAGGCGCTTCAAGTATGATGAGTGTCTTACCTATGGCGCTGC 2127
 QY 2059 AACTTCTCTGAGCTCCGCGCCATACCGCTACATACAGCTTACCTCCACCCCTGTAC 2118
 Db 2128 ACAGCCCTGAGAGCTCTTCCA----- 2150
 QY 2119 TAGGCGCTTACCTCCACCCCTGTCTACTACAGCTTACCTCTACAGGCTTAAGTCCAG 2178
 Db 2151 -----GCTGCCCTCAGACCTTTCGACCTTCTGAGGCTTAAGTCTAG 2194
 QY 2179 GCCATGCTGTGCTTCCAGAGGCTTCAAGACTTCTTAACCTGGAGTGTAGAGGAGC-T 2237
 Db 2195 CCCATATCTACTTGTAAAGGCTCAATCTTTTCTACATAGAGTGTGGTTAAGCGGT 2254
 QY 2238 GAAGTACCTTTGGGGCAACAGACACCTTATGTTTCAATCTGACCTGTAGCCCTGCACT 2297
 Db 2255 TTAAGGTGCTTTAGGATATACACCTAGTTTATTTGATGACCTACCTATTAACATT 2314
 QY 2298 CACCTGTGACAGCAATGAACAGAGCTTCCGTCGCAAAAAGGTCACGCTCCACCC 2357
 Db 2315 GTGTACAGGCTGAGATTAAACAGAAATTTTAAAGGTGGGTTTCTAGTTCAATGTC 2374
 QY 2358 CCGCCGCT-----CCCTGACCTCTCTGCTCTCTCTCCAGTTATTC 2399
 Db 2375 CCTTGCCCTGAGCTTGTGGCATTTCCACCTGCTGCTTCCACCTCTCTCTCTCTGAGTCC 2434
 QY 2400 TGAACACGCGAGGAGGAGCAACAGTGGCGCCCAAGAGGAGGAGGAGGAGGAGGAGG 2459
 Db 2435 TGAAGACAGCTAGGCGAGGAGGAGTGTG-CCACAAAGGAGATGACATTTTCAGTCCG 2493
 QY 2460 AGCGCGGAGAGGCTGGAAGGCGGAGATGCTGCTCCATCCATCCAGCTCAGGCTCAG 2519
 Db 2494 AGTCCACAGGCTGTAGGGGGCTCAGAGTCAACCCAGCCACCCAGCTACCCG-----A 2548
 QY 2520 GTCTTGTGCTGTCCCGAGACCTCTGTGAACCAAGCAGATCAAGGAGGAGGAGGAGG 2579
 Db 2549 ATCTGTGCTGTGCGCTGAGTTCCCGATGCC-----TCACTGGTCAAGCAG 2599
 QY 2580 AGAGATAGTCTTTTGTCTCTTCTGCGCTGTGCTAGTGTG-AGTTTTTCAATAGCTTA 2638
 Db 2600 TGAGA-----GTCTTCTTCTTCTTGAAGATCTCTGCGCAGGAGAGTTTTCATAGTGA 2654
 QY 2639 CAGTATGCGCTTGTACTGAGAAATAAACATTTTATTAATAAAAAA 2690
 Db 2655 TAGTATTGGCTTTGTATTGAGAAATAAAGACATTTTATTAATAAGCA 2706
 RESULT 15
 AP002776 160903 bp DNA linear HTG 17-JAN-2001
 LOCUS Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT
 DEFINITION
 ACCESSION AP002776
 VERSION AP002776.2 GI:12246853
 HTG: HTGS, PHASE1: HTGS, DRAFT.
 KEYWORDS Homo sapiens DNA, clone:RP11-126P21.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 160,903 genomic DNA of 11q
 Published Only in DataBase (2000)
 2 (bases 1 to 160903)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (10-JUL-2000) Masahisa Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gs.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/

COMMENT

```
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 16, 2001 this sequence version replaced gi:9188614.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hatori@gsc.riken.go.jp
----- Project Information
Center project name: Humpdraft1
Center clone name: Rp11-126p21
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 14950 bases at least Q40
Consensus quality: 15393 bases at least Q30
Consensus quality: 15614 bases at least Q20
Insert size: 157403; sum-of-contigs
Quality coverage: 8.42x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
36 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

1      12626 contig of 12626 bp in length
12727      24539 contig of 11813 bp in length
24640      33605 contig of 8966 bp in length
33706      45529 contig of 11824 bp in length
45630      51226 contig of 5597 bp in length
51327      58567 contig of 7241 bp in length
58668      65488 contig of 6821 bp in length
65589      72315 contig of 6727 bp in length
72416      78297 contig of 5882 bp in length
78398      84113 contig of 5716 bp in length
84214      90536 contig of 6323 bp in length
90537      90636 contig of 100 bp
90637      95051 contig of 4415 bp in length
95152      95151 contig of 100 bp
95152      99283 contig of 4132 bp in length
99284      99383 contig of 100 bp
99384      103658 contig of 4275 bp in length
103659     103758 contig of 100 bp
103759     107823 contig of 4065 bp in length
107824     107923 contig of 100 bp
107924     113135 contig of 5312 bp in length
113136     113235 contig of 100 bp
113236     117634 contig of 4399 bp in length
117635     117734 contig of 100 bp
117735     121241 contig of 3507 bp in length
121242     121341 contig of 100 bp
121342     124894 contig of 3553 bp in length
124895     124994 contig of 100 bp
124995     128257 contig of 3263 bp in length
128258     128357 contig of 100 bp
128358     131408 contig of 3051 bp in length
131409     131508 contig of 100 bp
131509     134212 contig of 2704 bp in length
134213     134312 contig of 100 bp
134313     136588 contig of 2276 bp in length
136589     136688 contig of 100 bp
136689     139497 contig of 2809 bp in length
139498     139597 contig of 100 bp
139598     142074 contig of 2477 bp in length
142075     142174 contig of 100 bp
142175     143858 contig of 1684 bp in length
143859     143958 contig of 100 bp
143959     146242 contig of 2284 bp in length
146243     146342 contig of 100 bp
146343     148167 contig of 1825 bp in length
148168     148267 contig of 100 bp
148268     150246 contig of 1979 bp in length
150247     150346 contig of 100 bp
150347     151876 contig of 1530 bp in length
151877     151976 contig of 100 bp
151977     154322 contig of 2246 bp in length
154323     154422 contig of 100 bp
154423     155669 contig of 1247 bp in length
155670     155769 contig of 100 bp
155770     156932 contig of 1163 bp in length
156933     157032 contig of 100 bp
157033     158659 contig of 1627 bp in length
158660     158759 contig of 100 bp
158760     159622 contig of 863 bp in length
159623     159722 contig of 100 bp
159723     160903 contig of 1181 bp in length.

* 12727      24539: contig of 11813 bp in length
* 24540      24639: gap of 100 bp
* 24640      33605: contig of 8966 bp in length
* 33606      33705: gap of 100 bp
* 33706      45529: contig of 11824 bp in length
* 45530      45629: gap of 100 bp
* 45630      51226: contig of 5597 bp in length
* 51227      51326: gap of 100 bp
* 51327      58567: contig of 7241 bp in length
* 58568      58667: gap of 100 bp
* 58668      65488: contig of 6821 bp in length
* 65489      65588: gap of 100 bp
* 65589      72315: contig of 6727 bp in length
* 72316      72415: gap of 100 bp
* 72416      78297: contig of 5882 bp in length
* 78298      78397: gap of 100 bp
* 78398      84113: contig of 5716 bp in length
* 84114      84213: gap of 100 bp
* 84214      90536: contig of 6323 bp in length
* 90537      90636: gap of 100 bp
* 90637      95051: contig of 4415 bp in length
* 95052      95151: gap of 100 bp
* 95152      99283: contig of 4132 bp in length
* 99284      99383: gap of 100 bp
* 99384      103658: contig of 4275 bp in length
* 103659     103758: gap of 100 bp
* 103759     107823: contig of 4065 bp in length
* 107824     107923: gap of 100 bp
* 107924     113135: contig of 5312 bp in length
* 113136     113235: gap of 100 bp
* 113236     117634: contig of 4399 bp in length
* 117635     117734: gap of 100 bp
* 117735     121241: contig of 3507 bp in length
* 121242     121341: gap of 100 bp
* 121342     124894: contig of 3553 bp in length
* 124895     124994: gap of 100 bp
* 124995     128257: contig of 3263 bp in length
* 128258     128357: gap of 100 bp
* 128358     131408: contig of 3051 bp in length
* 131409     131508: gap of 100 bp
* 131509     134212: contig of 2704 bp in length
* 134213     134312: gap of 100 bp
* 134313     136588: contig of 2276 bp in length
* 136589     136688: gap of 100 bp
* 136689     139497: contig of 2809 bp in length
* 139498     139597: gap of 100 bp
* 139598     142074: contig of 2477 bp in length
* 142075     142174: gap of 100 bp
* 142175     143858: contig of 1684 bp in length
* 143859     143958: gap of 100 bp
* 143959     146242: contig of 2284 bp in length
* 146243     146342: gap of 100 bp
* 146343     148167: contig of 1825 bp in length
* 148168     148267: gap of 100 bp
* 148268     150246: contig of 1979 bp in length
* 150247     150346: gap of 100 bp
* 150347     151876: contig of 1530 bp in length
* 151877     151976: gap of 100 bp
* 151977     154322: contig of 2246 bp in length
* 154323     154422: gap of 100 bp
* 154423     155669: contig of 1247 bp in length
* 155670     155769: gap of 100 bp
* 155770     156932: contig of 1163 bp in length
* 156933     157032: gap of 100 bp
* 157033     158659: contig of 1627 bp in length
* 158660     158759: gap of 100 bp
* 158760     159622: contig of 863 bp in length
* 159623     159722: gap of 100 bp
* 159723     160903: contig of 1181 bp in length.

FEATURES
Source
1. 160903
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

```
/chromosome="11"
/map="11q"
/clone="Rp11-126p21"
1..12626
misc_feature
    /note="assembly_fragment"
    12727..24539
misc_feature
    /note="assembly_fragment"
    24640..33605
misc_feature
    /note="assembly_fragment"
    33706..45529
misc_feature
    /note="assembly_fragment"
    45630..51226
misc_feature
    /note="assembly_fragment_clone_end:sp6_vector_side:left"
    51327..58567
misc_feature
    /note="assembly_fragment"
    58668..65488
misc_feature
    /note="assembly_fragment"
    65589..72315
misc_feature
    /note="assembly_fragment"
    72416..78297
misc_feature
    /note="assembly_fragment"
    78398..84113
misc_feature
    /note="assembly_fragment"
    84214..90536
misc_feature
    /note="assembly_fragment"
    90637..95051
misc_feature
    /note="assembly_fragment"
    95152..99283
misc_feature
    /note="assembly_fragment"
    99384..103658
misc_feature
    /note="assembly_fragment"
```

Query Match 37.1%; Score 1004; DB 2; Length 160903;

Best Local Similarity 100.0%; Pred. No. 1e-197;

Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1677 AGGCTCTCTCTCCAGAGTCTTCACATGAAGAGCCTCTGCAGCCCTTCCACAGCTTG 1736
    |||||||
DB 79592 AGGCTCTCTCTCCAGAGTCTTCACATGAAGAGCCTCTGCAGCCCTTCCACAGCTTG 79651

QY 1737 CAAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1796
    |||||||
DB 79652 CAAGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 79711

QY 1797 AGTCAGTGTGTTACCTCCAGGGAGTGGCTGTGGTGGCAACCGGACCCAGGCTTCCAGG 1856
    |||||||
DB 79712 AGTCAGTGTGTTACCTCCAGGGAGTGGCTGTGGTGGCAACCGGACCCAGGCTTCCAGG 79771

QY 1857 AGCAGAGCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1916
    |||||||
DB 79772 AGCAGAGCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 79831

QY 1917 TCCGGAAGGTGTGAGACAGGCGTGCATGACAGTGGAGAGGAGGAGGAGGAGGAGGAGG 1976
    |||||||
DB 79832 TCCGGAAGGTGTGAGACAGGCGTGCATGACAGTGGAGAGGAGGAGGAGGAGGAGGAGG 79891

QY 1977 CCTCAGACATGCTCCAGCGTCCCTGACACTGAAGAGATCCACAACCTCTTGAGAGAAC 2036
    |||||||
DB 79892 CCTCAGACATGCTCCAGCGTCCCTGACACTGAAGAGATCCACAACCTCTTGAGAGAAC 79951

QY 2037 ACCCTCAGCTGTGTGGCCGACACATTTCTCAGCTCGGCCCCATACCCGCTACATACAG 2096
    |||||||
DB 79952 ACCCTCAGCTGTGTGGCCGACACATTTCTCAGCTCGGCCCCATACCCGCTACATACAG 80011

QY 2097 CCTCAGCTTCCACCCCTGTGCTACAGGCTTCACTCCACCCCTGTGCTACAGCTTCAAC 2156
    |||||||
DB 80012 CCTCAGCTTCCACCCCTGTGCTACAGGCTTCACTCCACCCCTGTGCTACAGCTTCAAC 80071

QY 2157 CTCTCAGAGCTTTAAGTCCAGAGCCATGTGCTGCTGCTCAAGGGGCTCAAGACTTTCTAA 2216
    |||||||
DB 80072 CTCTCAGAGCTTTAAGTCCAGAGCCATGTGCTGCTGCTCAAGGGGCTCAAGACTTTCTAA 80131

QY 2217 CTGGGATGTGTAGAGGAGTGAAGGTACCTTTGGGGGCAACAGCACCTAGTTTCATTTC 2276
    |||||||
```

```
|||||
DB 80132 CTGGGATGTGTAGAGGAGTGAAGGTACCTTTGGGGGCAACAGCACCTAGTTTCATTTC 80191

QY 2277 TCACCTCTAGCCCTGCACACTACCTGTGTGGACAGGATGAAGAGAGGCTTCCCGTGCAA 2336
    |||||||
DB 80192 TCACCTCTAGCCCTGCACACTACCTGTGTGGACAGGATGAAGAGAGGCTTCCCGTGCAA 80251

QY 2337 AAAGGTCAGCGCTCCACACCCCGCCCTCCCTGACCTGCTGTCCTCTCCAGTTCAAT 2396
    |||||||
DB 80252 AAAGGTCAGCGCTCCACACCCCGCCCTCCCTGACCTGCTGTCCTCTCCAGTTCAAT 80311

QY 2397 TCCTGGAACCAAGCCAGGCGCAGGCAACAGTGGGCCCCAAAGGAGGAGGAGATCTCAGGC 2456
    |||||||
DB 80312 TCCTGGAACCAAGCCAGGCGCAGGCAACAGTGGGCCCCAAAGGAGGAGGAGATCTCAGGC 80371

QY 2457 CCCAGCGCGGGAGGCTGGAAGGCTTGGCAGATCGCTTCCCTCATATCCACTCCACCGGTC 2516
    |||||||
DB 80372 CCCAGCGCGGGAGGCTGGAAGGCTTGGCAGATCGCTTCCCTCATATCCACTCCACCGGTC 80431

QY 2517 CAGGCTTTGCTGCTGTCCCGAGACCTCCTGTGACACACGCGAGATCACAGGGCACAG 2576
    |||||||
DB 80432 CAGGCTTTGCTGCTGTCCCGAGACCTCCTGTGACACACGCGAGATCACAGGGCACAG 80491

QY 2577 GCCAGAGATAGTCTTCTTTTGTCTTCTTGTGGCTTGGCTGAGTCAATTTTCATAGCCT 2636
    |||||||
DB 80492 GCCAGAGATAGTCTTCTTTTGTCTTCTTGTGGCTTGGCTGAGTCAATTTTCATAGCCT 80551

QY 2637 TACAGTATCTGGCTTGTACTGAGAAATAAACACATTTTCATA 2680
    |||||||
DB 80552 TACAGTATCTGGCTTGTACTGAGAAATAAACACATTTTCATA 80595
    |||||||
```

Search completed: April 11, 2003, 19:58:11

Job time : 14091.8 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 00:05:14 ; Search time 394.978 Seconds
(without alignments)
6005.049 Million cell updates/sec

Title:	US-09-761-640-1
Perfect score:	2704
Sequence:	1 cgtcctctcgtctctcgg.....aaaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Published: Applications: NA: *

- 1: /cgn2_6/prodata/2/pubnpa/us07_PUBCOMB.seq: *
- 2: /cgn2_6/prodata/2/pubnpa/PCT_NEW_PUB.seq: *
- 3: /cgn2_6/prodata/2/pubnpa/US06_NEW_PUB.seq: *
- 4: /cgn2_6/prodata/2/pubnpa/US06_PUBCOMB.seq: *
- 5: /cgn2_6/prodata/2/pubnpa/US07_NEW_PUB.seq: *
- 6: /cgn2_6/prodata/2/pubnpa/PCTUS_PUBCOMB.seq: *
- 7: /cgn2_6/prodata/2/pubnpa/US08_NEW_PUB.seq: *
- 8: /cgn2_6/prodata/2/pubnpa/US08_PUBCOMB.seq: *
- 9: /cgn2_6/prodata/2/pubnpa/US09_NEW_PUB.seq: *
- 10: /cgn2_6/prodata/2/pubnpa/US09_PUBCOMB.seq: *
- 11: /cgn2_6/prodata/2/pubnpa/US10_NEW_PUB.seq: *
- 12: /cgn2_6/prodata/2/pubnpa/US10_PUBCOMB.seq: *
- 13: /cgn2_6/prodata/2/pubnpa/US60_NEW_PUB.seq: *
- 14: /cgn2_6/prodata/2/pubnpa/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	2704	100.0	2704	10	US-09-761-640-1	Sequence 1, Appl1
2	2704	100.0	2704	10	US-09-761-640-7	Sequence 7, Appl1
3	2558	94.6	2852	10	US-09-761-640-2	Sequence 2, Appl1
4	2246.4	83.1	2540	10	US-09-761-640-3	Sequence 3, Appl1
5	1774	63.6	1960	9	US-09-955-732-1	Sequence 1, Appl1
6	1416	52.4	1416	9	US-09-955-732-20	Sequence 20, Appl1
7	336.8	12.5	1949	10	US-09-775-925-1	Sequence 1, Appl1
8	313.4	11.6	1711	10	US-09-775-925-5	Sequence 5, Appl1
9	313.2	11.6	3314	10	US-09-969-708-398	Sequence 398, Appl1
10	208.6	7.7	1052	10	US-09-775-925-7	Sequence 7, Appl1
11	163.4	6.0	969	10	US-09-808-701-12	Sequence 12, Appl1
12	98.8	3.7	1685	12	US-10-002-600-51	Sequence 51, Appl1
13	89.2	3.3	2139	9	US-09-865-993-3	Sequence 3, Appl1
14	89.2	3.3	2479	9	US-09-865-993-10	Sequence 10, Appl1
15	88	3.3	1209	9	US-09-902-941-1918	Sequence 1918, Ap
16	88	3.3	1209	9	US-09-849-626-1918	Sequence 1918, Ap
17	88	3.3	1209	9	US-10-017-754-1818	Sequence 1818, Ap
18	88	3.3	1238	9	US-09-736-451-803	Sequence 803, App
19	88	3.3	1238	9	US-09-902-941-803	Sequence 803, App

20	88	3.3	1238	9	US-09-849-628-803	Sequence 803, App
21	88	3.3	1238	9	US-10-017-754-803	Sequence 803, App
22	88	3.3	2064	9	US-09-736-457-825	Sequence 825, App
23	88	3.3	2064	9	US-09-902-941-825	Sequence 825, App
24	88	3.3	2064	9	US-09-849-628-825	Sequence 825, App
25	88	3.3	2064	9	US-10-017-754-825	Sequence 825, App
26	88	3.3	2109	9	US-09-736-457-826	Sequence 826, App
27	88	3.3	2109	9	US-09-902-941-826	Sequence 826, App
28	88	3.3	2109	9	US-09-849-628-826	Sequence 826, App
29	88	3.3	2109	9	US-10-017-754-826	Sequence 826, App
30	88	3.3	2444	9	US-10-071-766-25	Sequence 29, Appl
31	87.6	3.2	2470	10	US-09-964-8244-555	Sequence 55, Ap
32	87.6	3.2	2470	10	US-08-880-107-310	Sequence 330, A
33	86.4	3.2	1619	9	US-09-736-947-801	Sequence 801, App
34	86.4	3.2	1619	9	US-09-902-941-801	Sequence 801, App
35	86.4	3.2	1619	9	US-09-849-628-801	Sequence 801, App
36	86.4	3.2	1637	9	US-10-017-754-801	Sequence 801, App
37	86.4	3.2	4637	9	US-09-736-457-804	Sequence 804, App
38	86.4	3.2	4637	9	US-09-902-941-804	Sequence 804, App
39	86.4	3.2	4637	9	US-09-849-628-804	Sequence 804, App
40	86.4	3.2	4637	9	US-10-017-754-804	Sequence 804, App
41	84	3.1	2203	10	US-09-880-107-3849	Sequence 3849, A
42	80.6	3.0	2436	10	US-09-917-8004-485	Sequence 485, Ap
43	73.6	2.7	1337	9	US-09-964-899-46	Sequence 46, Appl
44	72	2.7	1143	9	US-10-184-893-3	Sequence 3, Appl
45	72	2.7	2650	9	US-10-184-833-1	Sequence 1, Appl

```

ALIGNMENTS

RESULT 1
US-09-761-640-1
; Sequence 1, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-1

Query Match      100.0%; Score 2704; DB 10; Length 2704;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGTCTTCTGCTGCTCTGGGGTCCAGACCTCTCCGCGGGGTTGAGGGAAGGGCCGTGCC 60
        |||||||
DB       1  GCTCTTCTGCTGCTGCGGGTCCAGACTGTCCGCGGGGTTGAGGGAAGGGCCGTGCC 60

QY      61  CGGTCCAGACCAGAGTGCCTGCGGCTGGCTGCATGTGCGCTGTACAGTAGAGCGTTGC 120
        |||||||
DB       61  CGGTCCAGACCAGAGTGCCTGCGGCTGGCTGCATGTGCGCTGTACAGTAGAGCGTTGC 120

QY      121  CCCCCGGGCGAGGGCGCTCCACAGCCGCTGGGGCCCTTGAGACAGGGCGGTCCAGCGAAG 180
        |||||||
DB       121  CCCCCGGGCGAGGGCGCTCCACAGCCGCTGGGGCCCTTGAGACAGGGCGGTCCAGCGAAG 180

QY      181  AGTCACATCCAGCGAAGGACAGAGCTTTCCGGTGTCCGTGGGGGCTGTGCTCCGGAGCTCAG 240
        |||||||
DB       181  AGTCACATCCAGCGAAGGACAGAGCTTTCCGGTGTCCGTGGGGGCTGTGCTCCGGAGCTCAG 240

QY      241  GATGAGGGGGAATGATGATGACAGACAGGCCAGTTCTGAGCCACAGAGAAAGGCCCCG 300
        |||||||
DB       241  GATGAGGGGGAATGATGATGACAGACAGGCCAGTTCTGAGCCACAGAGAAAGGCCCCG 300

```

Db	241	GATGGAGGGGACATGATGATGACAGACAGGCCAATTCTGAECCTCAACAGAGAAGGCCCGC	300
OY	301	AGTGAAGAGAGAGCTCCACGGGGAGCCAGACAGACTTCGGGCAAGAGATCCAGAGTCCCGAG	360
Db	301	AGTGAAGAGAGAGCTCCACGGGGAGCCAGACAGACTTCGGGCAAGAGATCCAGAGTCCCGAG	360
OY	361	AAGCAGAGAGAGACAGAGCAGCAGCTGCACTCATATGTAACAGCTGTGAGGCGCGAGAT	420
Db	361	AAGCAGAGAGAGACAGAGCAGCAGCTGCACTCATATGTAACAGCTGTGAGGCGCGAGAT	420
OY	421	GACATCCGCGTGGCAGGCCAGCTGGAGGAGACCCGGGCTCCCGGGCTGCGTACCTGCTAG	480
Db	421	GACATCCGCGTGGCAGGCCAGCTGGAGGAGACCCGGGCTCCCGGGCTGCGTACCTGCTAG	480
OY	481	GTAAGTTTCTACACGAGAAAGAGAGGTTGAGCCAGATGAGACGGTCTCTCTGGGCGTG	540
Db	481	GTAAGTTTCTACACGAGAAAGAGAGGTTGAGCCAGATGAGACGGTCTCTCTGGGCGTG	540
OY	541	GATTTTCCTGACAGCAGCTCCCCAGCTGCACCTTGGGCGCTGTCTTGGCCCTCTGGAGT	600
Db	541	GATTTTCCTGACAGCAGCTCCCCAGCTGCACCTTGGGCGCTGTCTTGGCCCTCTGGAGT	600
OY	601	GACACCCAGAGTACTTATATGATGAGAGCGGGGGTGTAGGCTGACGTCGTGATGGGCAAGC	660
Db	601	GACACCCAGAGTACTTATATGATGAGAGCGGGGGTGTAGGCTGACGTCGTGATGGGCAAGC	660
OY	661	CGGATCTTCAGAGCCCATCTCCATCCAGACCATATGTGGGCCACATCCAGGTATTGACCAA	720
Db	661	CGGATCTTCAGAGCCCATCTCCATCCAGACCATATGTGGGCCACATCCAGGTATTGACCAA	720
OY	721	GCATGTGAGGCAAGCTTAAAGCAGCGGCGCTTGTACGGGGTGGCAGTGGCCCTCACCTGGGCC	780
Db	721	GCATGTGAGGCAAGCTTAAAGCAGCGGCGCTTGTACGGGGTGGCAGTGGCCCTCACCTGGGCC	780
OY	781	AGCGACTTACAGAGAGAGACTGAACTCCGAACAGAGCTGCCTCATATGATGAGAGCGGTATG	840
Db	781	AGCGACTTACAGAGAGAGACTGAACTCCGAACAGAGCTGCCTCATATGATGAGAGCGGTATG	840
OY	841	GCCGACCTGAGAGTCTGTGGGGCTTCCAGGCGCCAGAGCTTGGCGGGTCTCTCAAGAACAGGAG	900
Db	841	GCCGACCTGAGAGTCTGTGGGGCTTCCAGGCGCCAGAGCTTGGCGGGTCTCTCAAGAACAGGAG	900
OY	901	CAGATGAGCAGGCGATCCGTGCTGAGCTGTGAAACTGTTGATGTCACTGACCTTGGAG	960
Db	901	CAGATGAGCAGGCGATCCGTGCTGAGCTGTGAAACTGTTGATGTCACTGACCTTGGAG	960
OY	961	AGGTGACTTCCAAAGAGATCGGCGAGGCTCTGAGCTGGCGCTGGGGGCTCCGCTTCAG	1020
Db	961	AGGTGACTTCCAAAGAGATCGGCGAGGCTCTGAGCTGGCGCTGGGGGCTCCGCTTCAG	1020
OY	1021	CAGTACCTGATCTTCATCCGACCAACAGATGCTGTGCTGTGGTCACAGCGGAGCCGAGCC	1080
Db	1021	CAGTACCTGATCTTCATCCGACCAACAGATGCTGTGCTGTGGTCACAGCGGAGCCGAGCC	1080
OY	1081	TCCCGATCTTCCCCACCTCTACCTTGGGCTCAAGATGGAACGAGCAACCTTGGAGGAG	1140
Db	1081	TCCCGATCTTCCCCACCTCTACCTTGGGCTCAAGATGGAACGAGCAACCTTGGAGGAG	1140
OY	1141	CTTCCAGAGGAACAGGGTCAACCATCTTAAACATGTGCCCCGGGAGATTGACAACTTTAC	1200
Db	1141	CTTCCAGAGGAACAGGGTCAACCATCTTAAACATGTGCCCCGGGAGATTGACAACTTTAC	1200
OY	1201	CTGAGAGGCTTCACTTCAACCAATGTGCGGCTCTGGAGTATGAGAGTGGGCCAGCTGCTG	1260
Db	1201	CTGAGAGGCTTCACTTCAACCAATGTGCGGCTCTGGAGTATGAGAGTGGGCCAGCTGCTG	1260
OY	1261	CCGCACTTGAAGAGAGACGACCCGTTTCATTGAGGCTCAAGAGACGAGGGACCCACAGTG	1320
Db	1261	CCGCACTTGAAGAGAGACGACCCGTTTCATTGAGGCTCAAGAGACGAGGGACCCACAGTG	1320
OY	1321	CTGTGCTCACTGCAAGATGGGCGTACGGCGCTCAGCGGCGACAGTGTCTGGCTATGCCATG	1380
Db	1321	CTGTGCTCACTGCAAGATGGGCGTACGGCGCTCAGCGGCGACAGTGTCTGGCTATGCCATG	1380
OY	1321	CTGTGCTCACTGCAAGATGGGCGTACGGCGCTCAGCGGCGACAGTGTCTGGCTATGCCATG	1380
Db	1321	CTGTGCTCACTGCAAGATGGGCGTACGGCGCTCAGCGGCGACAGTGTCTGGCTATGCCATG	1380

QY	1381	AAGCAGTACGAATGCAGCCTGGAGCAGGCCCTGCGCCACGTGCAGAGAGCTCCGGCCATTC	1440
Db	1381	AAGCAGTACGAATGCAGCCTGGAGCAGGCCCTGCGCCACGTGCAGAGAGCTCCGGCCATTC	1440
QY	1441	GCCCGCCCAACCCCTGGCTTCTTGCGCCACGTGCAGATTTACCAGGGGACCTCGAGCGGC	1500
Db	1441	GCCCGCCCAACCCCTGGCTTCTTGCGCCACGTGCAGATTTACCAGGGGACCTCGAGCGGC	1500
QY	1501	AGACATGGAGGGGTGGGGAGAGAGAAGTTGTAGGCATGGAAGAAGCAGGCAACCCCC	1560
Db	1501	AGACATGGAGGGGTGGGGAGAGAGAAGTTGTAGGCATGGAAGAAGCAGGCAACCCCC	1560
QY	1561	GAAGAAGAAGCCTGGGCCACAGGCCACGTAATAACCTCCGAGGGGTGATGAGTCCATCAG	1620
Db	1561	GAAGAAGAAGCCTGGGCCACAGGCCACGTAATAACCTCCGAGGGGTGATGAGTCCATCAG	1620
QY	1621	TCTTTCTGGAGCCTCCTTGGAGCTGGAGAGCACTCTAGAGACCAGTGCATGGCAGAGGT	1680
Db	1621	TCTTTCTGGAGCCTCCTTGGAGCTGGAGAGCACTCTAGAGACCAGTGCATGGCAGAGGT	1680
QY	1681	CTTCTCTCTCCACAGAGCTCTTCACATGAAGAGCCTCTGACAGCCTTCTCCACAGCTTCCAG	1740
Db	1681	CTTCTCTCTCCACAGAGCTCTTCACATGAAGAGCCTCTGACAGCCTTCTCCACAGCTTCCAG	1740
QY	1741	GACCAAGGAGGAGCCACAGGTGGACAGGGGGCCTCAGCCTGCCCTGAATGCCGCCAGTC	1800
Db	1741	GACCAAGGAGGAGCCACAGGTGGACAGGGGGCCTCAGCCTGCCCTGAATGCCGCCAGTC	1800
QY	1801	AGTGGTTACCTTCAGAGGCAGTGCCTGGTGGCCAAACCGGACCCAGGCCCTTCCAGAGACA	1860
Db	1801	AGTGGTTACCTTCAGAGGCAGTGCCTGGTGGCCAAACCGGACCCAGGCCCTTCCAGAGACA	1860
QY	1861	GGAGCAGGGGCAGGGGAGGGGAGGGAGAGAGCCCTGATTTCTCTACAGGCCAGTTCGCG	1920
Db	1861	GGAGCAGGGGCAGGGGAGGGGAGGGAGAGAGCCCTGATTTCTCTACAGGCCAGTTCGCG	1920
QY	1921	GGAAGTGTGTAGAGCAGGCCAGCGCTGCATGACAGTGAAGAGGAGGGCGAGGCCCTGAGCCCT	1980
Db	1921	GGAAGTGTGTAGAGCAGGCCAGCGCTGCATGACAGTGAAGAGGAGGGCGAGGCCCTGAGCCCT	1980
QY	1981	CACACATGCCACAGCCTCCCTGACACTGAAGAGAGATTCACAACTCCTTGGAGAAACACCC	2040
Db	1981	CACACATGCCACAGCCTCCCTGACACTGAAGAGAGATTCACAACTCCTTGGAGAAACACCC	2040
QY	2041	TCAAGTGTGTGGCGACACATTCCTCTCAAGTCGGGCCCATACCGCTGACTCAAGCCTC	2100
Db	2041	TCAAGTGTGTGGCGACACATTCCTCTCAAGTCGGGCCCATACCGCTGACTCAAGCCTC	2100
QY	2101	ACCTGCCACCCCTGTCACTACGGCTACCTCCACCCCTGTCACTACAGCCTCACCTCC	2160
Db	2101	ACCTGCCACCCCTGTCACTACGGCTACCTCCACCCCTGTCACTACAGCCTCACCTCC	2160
QY	2161	TACAGCCTTAAGTCCAGGCCCATGTCTGCCCTGTCCAAAGGCTTCAGACATTTCTAACTGG	2220
Db	2161	TACAGCCTTAAGTCCAGGCCCATGTCTGCCCTGTCCAAAGGCTTCAGACATTTCTAACTGG	2220
QY	2221	GATGTGGTAGAGGAGACGTGAAGAGTACCTTTGGGGGCAACAGACCTCGATTTCATTCGAA	2280
Db	2221	GATGTGGTAGAGGAGACGTGAAGAGTACCTTTGGGGGCAACAGACCTCGATTTCATTCGAA	2280
QY	2281	CTTAGGCCCTGCACACTACCTGTGGCAGCGAATGAAGAACAGAGCTTCGCCGTGCAAAAG	2340
Db	2281	CTTAGGCCCTGCACACTACCTGTGGCAGCGAATGAAGAACAGAGCTTCGCCGTGCAAAAG	2340
QY	2341	GGTAGCAGCTTCCACCCCGCCCTCCCTGACACTCTGTCTCTCTCCAGTTCATTCT	2400
Db	2341	GGTAGCAGCTTCCACCCCGCCCTCCCTGACACTCTGTCTCTCTCCAGTTCATTCT	2400
QY	2401	GGAACCAAGCAGGCCAGGCAACAGCTGGCCCCCAAGGCAAGGAGATCTCCAGGCCCA	2460
Db	2401	GGAACCAAGCAGGCCAGGCAACAGCTGGCCCCCAAGGCAAGGAGATCTCCAGGCCCA	2460

|||||
Db 1501 AGAAGCTAGGGTGTGGGAGAGAGAGTTGTAGCATGGAAGAGAGCCAGGAGCCCC 1560
QY 1561 GAAAGAGGCGCTGGGCGACGGCCACGTATAACCTCCGAGGGGCGATGAGGTCATCAG 1620
Db 1561 GAAAGAGGCGCTGGGCGACGGCCACGTATAACCTCCGAGGGGCGATGAGGTCATCAG 1620
QY 1621 TCTTCTGGAGCCCTCTTGTGAGCTGGAGAGCACCTCAGAGACAGTGAATCCAGAGGT 1680
Db 1621 TCTTCTGGAGCCCTCTTGTGAGCTGGAGAGCACCTCAGAGACAGTGAATCCAGAGGT 1680
QY 1681 CTTTCTTCTCCACGAGTCTTCAATGAGAGGCTCTGCAGCCCTTCCACACTTGGCAAG 1740
Db 1681 CTTTCTTCTCCACGAGTCTTCAATGAGAGGCTCTGCAGCCCTTCCACACTTGGCAAG 1740
QY 1741 GAGCAAGGAGGCGCAGAGTGGACAGAGGGGGCTCAGCTGCGCTGGAAGTCCGCGCAGTC 1800
Db 1741 GAGCAAGGAGGCGCAGAGTGGACAGAGGGGGCTCAGCTGCGCTGGAAGTCCGCGCAGTC 1800
QY 1801 AGTGGTTACCTTCAGAGGAGTGCCTGGTGGCCAAACCGGAGCCAGCCCTTCAGAGAGA 1860
Db 1801 AGTGGTTACCTTCAGAGGAGTGCCTGGTGGCCAAACCGGAGCCAGCCCTTCAGAGAGA 1860
QY 1861 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGGCCCTGCAATTTCTTACGCCAGGTTCCG 1920
Db 1861 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGGCCCTGCAATTTCTTACGCCAGGTTCCG 1920
QY 1921 GAAGGTGTGAGACAGAGGCGAGGCTGATAGAGTGAAGAGAGGGGCGAGGCGCTGAGCCCT 1980
Db 1921 GAAGGTGTGAGACAGAGGCGAGGCTGATAGAGTGAAGAGAGGGGCGAGGCGCTGAGCCCT 1980
QY 1981 CACACATGCCCAAGCGTCCCTGACACTGAAGAGATCCAACTCCTTGGAGAAACACC 2040
Db 1981 CACACATGCCCAAGCGTCCCTGACACTGAAGAGATCCAACTCCTTGGAGAAACACC 2040
QY 2041 TCACGCTTGTGGCGACACATTCCTCTCAGCTCCGCGCCCATACCCGTACTACAGACCTC 2100
Db 2041 TCACGCTTGTGGCGACACATTCCTCTCAGCTCCGCGCCCATACCCGTACTACAGACCTC 2100
QY 2101 ACCTCCACCCCTGTGACATACAGGCGCTCACCTCCACCCCTGTACTACAGGCTCACCTCC 2160
Db 2101 ACCTCCACCCCTGTGACATACAGGCGCTCACCTCCACCCCTGTACTACAGGCTCACCTCC 2160
QY 2161 TACAGCCTTAAGTCCAGGCGCATGTCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGG 2220
Db 2161 TACAGCCTTAAGTCCAGGCGCATGTCTGCTGCTCCAGAGGCTCAAGACTTTTAACTGG 2220
QY 2221 GATGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCTAGTTTCTCA 2280
Db 2221 GATGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCTAGTTTCTCA 2280
QY 2281 CTTTACCCCTTGACACTCACTGTGGCAGGAAATGAAGAGAGCTTCCGTCAGAAAAG 2340
Db 2281 CTTTACCCCTTGACACTCACTGTGGCAGGAAATGAAGAGAGCTTCCGTCAGAAAAG 2340
QY 2341 GGTACAGCCTCCACCCCGCCCTTCCCTGACACTCTCTCTCTCCAGTTTCAATTCCT 2400
Db 2341 GGTACAGCCTCCACCCCGCCCTTCCCTGACACTCTCTCTCTCCAGTTTCAATTCCT 2400
QY 2401 GGAACACGAGGCGAGGCAACAGAGTGGCCCGCAAGAGCGAGAGATCTCGAGCCCA 2460
Db 2401 GGAACACGAGGCGAGGCAACAGAGTGGCCCGCAAGAGCGAGAGATCTCGAGCCCA 2460
QY 2461 GCGGCGGAGAGCTGGAAGGCTGGAGATGCTTCCCTCATTCACCTCCACGCGTCCAGG 2520
Db 2461 GCGGCGGAGAGCTGGAAGGCTGGAGATGCTTCCCTCATTCACCTCCACGCGTCCAGG 2520
QY 2521 TCTTGTGCTGTCTCCAGACTCTCTGTGACACACGAGATCAGAGGCGACAGGCCA 2580
Db 2521 TCTTGTGCTGTCTCCAGACTCTCTGTGACACACGAGATCAGAGGCGACAGGCCA 2580
QY 2581 GAATATGCTCTTCTTGTGCTTCTGCGCTCTGAGTCAATTTTCACTAACCTTACA 2640
|||||

Db 2581 GAGATAGTCTTCTTGTGCTTCTGCGCTCTGAGTCAATTTTCACTAACCTTACA 2640
QY 2641 GATATGCGCTTGTACTAGAAATATAACATTTTCATTAATAAAAAAAAAAAAAA 2700
Db 2641 GATATGCGCTTGTACTAGAAATATAACATTTTCATTAATAAAAAAAAAAAAAA 2700
QY 2701 AAAA 2704
Db 2701 AAAA 2704

RESULT 3
US-09-761-640-2
; Sequence 2, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2852
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-2

Query Match 94.6% Score 2558: DB 10: Length 2852;
Best Local Similarity 96.5% Pred. No. 0;
Matches 2666: Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY 39 GGTGAGGAGAGGGGCGTGGCCGAGTGCAGCCAGTGTCTGCGGCGTGCCTCCATGCG 98
Db 2 GGTGAGGAGAGGGGCGTGGCCGAGTGCAGCCAGTGTCTGCGGCGTGCCTCCATGCG 61

QY 99 CCTGTGACAGTGAAGCGTTTGGCCCCCGGCGAGCGGCGCTTCCAGCGCCGTTGGGCGCTG 158
Db 62 CCGTGTACAGTGAAGCGTTTGGCCCCCGGCGAGCGGCGCTTCCAGCGCCGTTGGGCGCTG 121

QY 159 GGAACGAGGGGCTCCAGGAGAGTCACTCAGGAGAGGAGAGGAGGCTTGGGCTGCG 218
Db 122 GGAACGAGGGGCTCCAGGAGAGTCACTCAGGAGAGGAGAGGAGGCTTGGGCTGCG 181

QY 219 TGGGCGTGTCTGGGACTCAGAGATGAGAGGAGCAATGATGATGAGAGCAGAGGCCAGTTTC 278
Db 182 TGGGCGTGTCTGGGACTCAGAGATGAGAGGAGCAATGATGATGAGAGCAGAGGCCAGTTTC 241

QY 279 TGAAGCCAGAGAGAGGCCCCGAGATGAGAGAGCTCCACGGGAGACAGACACTTGG 338
Db 242 TGAAGCCAGAGAGAGGCCCCGAGATGAGAGAGCTCCACGGGAGACAGACACTTGG 301

QY 339 GCAAGATATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGTCCACCTCATGTT 398
Db 302 GCAAGATATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGTCCACCTCATGTT 361

QY 399 ACAGTGTGAGGCGCAGAGATGACATCCGCTGGAGGCCAGCTGAGAGCACCCCGGCC 458
Db 362 ACAGTGTGAGGCGCAGAGATGACATCCGCTGGAGGCCAGCTGAGAGCACCCCGGCC 421

QY 459 TCCCGGCTCCGCTACCTCTGGTACTTCTACAGAGAGAGAGAGAGTCTAGGCCAGGA 518
Db 422 TCCCGGCTCCGCTACCTCTGGTACTTCTACAGAGAGAGAGAGAGTCTAGGCCAGGA 481

QY 519 TGAAGAGTCTCTCTGGGCGTGAATTTCCCTGACAGACACTCCCAAGTGCACCTGGG 578
Db 482 TGAAGAGTCTCTCTGGGCGTGAATTTCCCTGACAGACACTCCCAAGTGCACCTGGG 541

QY 579 CCGGCTTGGCCCTGTGGAGTACACCCAGGCTGACTTGAATGAGAGAGGGGCGCTTCA 638
|||||

Db 542 CCTGCTTTGCCCTCTGTGAGTGACACCAGGTACTTAGATGAGAGCGGGGCTTCAG 601
Qy 639 CGTGCAGCTGTGGTGGGCAAAAGCCGGATCTTCAAGCCCATCTTCATCCAGACCATGTMGGGC 698
Db 602 CGTGACGCTGTGGTGGGCAAAAGCCGGATCTTCAAGCCCATCTTCATCCAGACCATGTMGGGC 661
Qy 699 CACACATCCAGGATATGACACCAAGCATGTGAGGACAGCTCTAGGACAGCGGCTTTGACCGGG 758
Db 662 CACACTCCAGGATATGACACCAAGCATGTGAGGACAGCTCTAGGACAGCGGCTTTGACCGGG 721
Qy 759 TGGCAGTGCCTCACCCTGGGCGGACGCACTACAGAGAGAGACTGAACCTCCGAACAGAGCTG 818
Db 722 TGGCAGTGCCTCACCCTGGGCGGACGCACTACAGAGAGAGACTGAACCTCCGAACAGAGCTG 781
Qy 819 CCTCAATGATGAGGAGGATGATGAGCGGACCTGGAGTCTCTGGCGGCTCCAGAGCGGAGCC 878
Db 782 CCTCAATGATGAGGAGGATGATGAGCGGACCTGGAGTCTCTGGCGGCTCCAGAGCGGAGCC 841
Qy 879 TGGCGGCTCTCAGAACAGAGAGAGATGAGAGCGGATCCGTGCTGAGCTGTGAAAGT 938
Db 842 TGGCGGCTCTCAGAACAGAGAGAGATGAGAGCGGATCCGTGCTGAGCTGTGAAAGT 901
Qy 939 GTTGGATGTCAGTGCAGTGGAGAGTGTCACTTCCAAAGAGATCCCGCCAGGCTCTGGAGCT 998
Db 902 GTTGGATGTCAGTGCAGTGGAGAGTGTCACTTCCAAAGAGATCCCGCCAGGCTCTGGAGCT 961
Qy 999 GCGCCTGGGGCTCCCGCTCCAGAGATACCGTGCATTCATGCAACACAGATGCTGCTGCT 1058
Db 962 GCGCCTGGGGCTCCCGCTCCAGAGATACCGTGCATTCATGCAACACAGATGCTGCTGCT 1021
Qy 1059 GGTGACACAGCGGGAGCGAGCCTCCGCACTTCCCGCACTCTACCTGGGCTGACAGAGT 1118
Db 1022 GGTGACACAGCGGGAGCGAGCCTCCGCACTTCCCGCACTCTACCTGGGCTGACAGAGT 1081
Qy 1119 GAAGCAGCAAAACCTGGAGAGAGCTCAGAGAGAACAGGGTCAACCCACATCTTAAATAGGC 1178
Db 1082 GAAGCAGCAAAACCTGGAGAGAGCTCAGAGAGAACAGGGTCAACCCACATCTTAAATAGGC 1141
Qy 1179 CCGGAGAGATTGACAACCTTCACTCCCTGAGAGCGCTTCACTACCAATGTGCGGCTTGCGGA 1238
Db 1142 CCGGAGAGATTGACAACCTTCACTCCCTGAGAGCGCTTCACTACCAATGTGCGGCTTGCGGA 1201
Qy 1239 TGAGAGTTCGGCCAGCTGCTGCGCACTGGAAGAGAGACGACCGCTTCATTTGAGGCTGC 1298
Db 1202 TGAGAGTTCGGCCAGCTGCTGCGCACTGGAAGAGAGACGACCGCTTCATTTGAGGCTGC 1261
Qy 1299 AAGACACAGGGGACACCGAGTGTGCTGCTGCACTGCAAGATGGGCGCTGACGCGGCTGC 1358
Db 1262 AAGACACAGGGGACACCGAGTGTGCTGCTGCACTGCAAGATGGGCGCTGACGCGGCTGC 1321
Qy 1359 CACAGTGTGAGCTATGSCATGTAAGACATAGCAATGCAAGCTGAGAGAGCGGCTCGGCCA 1418
Db 1322 CACAGTGTGAGCTATGSCATGTAAGACATAGCAATGCAAGCTGAGAGAGCGGCTCGGCCA 1381
Qy 1419 CGTGAGAGAGCTCGGGCCCATGCGCGCCCAACCTTGCTTCTGCGCCAGCTGCAAGT 1478
Db 1382 CGTGAGAGAGCTCGGGCCCATGCGCGCCCAACCTTGCTTCTGCGCCAGCTGCAAGT 1441
Qy 1479 CTACCAGGGGCAATCTGAGG----- 1497
Db 1442 CTACCAGGGGCAATCTGAGGCGCCAGCGCCAGAGCCATGCTCTGGAGAGAGAAAGTGGGTGG 1501
Qy 1498 -----GCC 1500
Db 1502 GGTCTCCCGCAGAGAGAGCAACCGCCCTGAAGTCTCTACACATTCCCACTTCTTCGCGC 1561
Qy 1501 AGAAGCTTAGGGTGTGGGAGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGAGCGCCC 1560
Db 1562 AGAAGCTTAGGGTGTGGGAGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGAGCGCCC 1621
Qy 1561 GAAAGAAAGAGCTTGGGCCACAGCGCCAGCTATAAAGCTCCGAGGGGCTCATAGAGTCCATCAG 1620
Db 1622 GAAAGAAAGAGCTTGGGCCACAGCGCCAGCTATAAAGCTCCGAGGGGCTCATAGAGTCCATCAG 1681

Qy 1621 TCTTCTGGAGCCCTCCTTGGAGCTGGAGAGACACTCAGAGACCAAGTACATGACATGACAGAGT 1680
Db 1682 TCTTCTGGAGCCCTCCTTGGAGCTGGAGAGACACTCAGAGACCAAGTACATGACATGACAGAGT 1741
Qy 1681 CTTCCTTCCCAAGAGTCTTCACATGAAGAGGCTCTGCAAGCCCTTCCCAAGCTTGGCAAG 1740
Db 1742 CTTCCTTCCCAAGAGTCTTCACATGAAGAGGCTCTGCAAGCCCTTCCCAAGCTTGGCAAG 1801
Qy 1741 GACCAAGGAGGCCACAGAGTGTGACAGAGGGGCTCAGGCTGCGCTGGAAGTCCCGGCAGTGC 1800
Db 1802 GACCAAGGAGGCCACAGAGTGTGACAGAGGGGCTCAGGCTGCGCTGGAAGTCCCGGCAGTGC 1861
Qy 1801 AGTGCTTACCTCCAGAGGAGTGTGCTGAGCAACCCAGGACCCAGGCTTCCAGAGAGA 1860
Db 1862 AGTGCTTACCTCCAGAGGAGTGTGCTGAGCAACCCAGGACCCAGGCTTCCAGAGAGA 1921
Qy 1861 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGCCGTGCATTTCTCTACGCCAGGTTCCG 1920
Db 1922 GGAGCAGGGGCGAGGGGCGAGGGGAGAGAGCCGTGCATTTCTCTACGCCAGGTTCCG 1981
Qy 1921 GAAGGTGGTGAACAGGGCCAGGCTGCATGACATGAGAGAGAGGGGCGAGGCTTGAGCCCT 1980
Db 1982 GAAGGTGGTGAACAGGGCCAGGCTGCATGACATGAGAGAGAGGGGCGAGGCTTGAGCCCT 2041
Qy 1981 CACACATGCGCCACAGCTCCCTGACACTGAGAGAGATCCACACACTCTTGGAGAGAACACC 2040
Db 2042 CACACATGCGCCACAGCTCCCTGACACTGAGAGAGATCCACACACTCTTGGAGAGAACACC 2101
Qy 2041 TCAGCTCTGTGTCGACACATTCCTCTCAGTCCGCCCCCATACCCGTCAACAGCCTC 2100
Db 2102 TCAGCTCTGTGTCGACACATTCCTCTCAGTCCGCCCCCATACCCGTCAACAGCCTC 2161
Qy 2101 ACCTGCCACCCCTGTGCACAGGCGCTCACTACCTCCACCCCTGTACATACAGCCTCACTCC 2160
Db 2162 ACCTGCCACCCCTGTGCACAGGCGCTCACTACCTCCACCCCTGTACATACAGCCTCACTCC 2221
Qy 2161 TACAGCCTTAAAGTCCACAGGCCATGTGCTGCTGCAAGAGGCTCAAGACTTCTTAACTGG 2220
Db 2222 TACAGCCTTAAAGTCCACAGGCCATGTGCTGCTGCAAGAGGCTCAAGACTTCTTAACTGG 2281
Qy 2221 GATGTGTAGAGGAGCTGAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCAATTCGAA 2280
Db 2282 GATGTGTAGAGGAGCTGAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCAATTCGAA 2341
Qy 2281 CTCTAGCCCTGACACACTCACTGTGGACAGGAATGAAGAAAGAGCTTCCCGTGAAGAAAG 2340
Db 2342 CTCTAGCCCTGACACACTCACTGTGGACAGGAATGAAGAAAGAGCTTCCCGTGAAGAAAG 2401
Qy 2341 GGTACAGGCTCCACACCCCGCCCTGCACTGCTGCTGCTCCAGATTCATTCCT 2400
Db 2402 GGTACAGGCTCCACACCCCGCCCTGCACTGCTGCTGCTCCAGATTCATTCCT 2461
Qy 2401 GGAACACAGCCAGGCCACAGCACTGAGGCCCCAAGAGGAGGAGAGATCTTCAGGCCCA 2460
Db 2462 GGAACACAGCCAGGCCACAGCACTGAGGCCCCAAGAGGAGGAGAGATCTTCAGGCCCA 2521
Qy 2461 GCGCGGGAGAGGCTGGAAGGGCTGGAGATGCTTCCCTATCCAGCTCCACCGGCTCAGG 2520
Db 2522 GCGCGGGAGAGGCTGGAAGGGCTGGAGATGCTTCCCTATCCAGCTCCACCGGCTCAGG 2581
Qy 2521 TCTTTGCTGCTGCCACAGACTCTGTGACACACAGGCAATCAGAGGCAACGAGCCA 2580
Db 2582 TCTTTGCTGCTGCCACAGACTCTGTGACACACAGGCAATCAGAGGCAACGAGGCA 2641
Qy 2581 GAGATAGTCTTTCTTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTTCAATAGCCTTACA 2640
Db 2642 GAGATAGTCTTTCTTTTGTCTTCTGCTGCTGCTGCTGCTGCTGCTTCAATAGCCTTACA 2701
Qy 2641 GTATCTGGCTTTGTACTGAGAAATTAACACATTTTTCATTAACAAAAA 2700
Db 2702 GTATCTGGCTTTGTACTGAGAAATTAACACATTTTTCATTAACAAAAA 2761

Oy 2701 AAAA 2704
|||||
Db 2762 AAAA 2765

RESULT 4

US-09-761-640-3
; Sequence 3, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: CLO00964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2540
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-3

Query Match 83.1%; Score 2246.4; DB 10; Length 2540;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5;

Oy 8 CCTGTCTCGCGGTCACAGACTGT-CCGCGGGGTTGAGGGAAGGGCCGTCGCCGTC 66
Db 1 CCGTGTCTCGCGGTCACAGACTGTCCCGGGGTTGAGGGAAGGGCCGTCGCCGTC 60
Oy 67 CAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGTCACAGTGAAGCCCTTCCGCCCG 126
Db 61 CAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGTCACAGTGAAGCCCTTCCGCCCG 120
Oy 127 GCGACGGGCGCTCCACGCGCCCTGGGGCCCTGGAGCAGGCGGTCCAGCGAAGACTGA 186
Db 121 GCGACGGGCGCTCCACGCGCCCTGGGGCCCTGGAGCAGGCGGTCCAGCGAAGACTGA 180
Oy 187 CTCACAGGAAGCAGAGCTTTGGGTGCTCCGTGGGGGCTCTCTGGGACTGCAGATGA 246
Db 181 CTCACAGGAAGCAGAGCTTTGGGTGCTCCGTGGGGGCTCTCTGGGACTGCAGATGA 240
Oy 247 GGGGACATGATGATCAGCAGAGGCGCAGTTCAGACCAAGAGAGAGGCCCGAGATGAG 306
Db 241 GGGGACATGATGATCAGCAGAGGCGCAGTTCAGACCAAGAGAGAGGCCCGAGATGAG 300
Oy 307 GAGGAGCTCCAGGGGAGCAGACAGACTTTCGGGCAAGATCCAGAGTCCCGCAGAAACAG 366
Db 301 GAGGAGCTCCAGGGGAGCAGACAGACTTTCGGGCAAGATCCAGAGTCCCGCAGAAACAG 360
Oy 367 GAGGAGCAAGAGCAGACCTGCACTCATGATGATGAGTCTGTGAGGGCGGAGATGATC 426
Db 361 GAGGAGCAAGAGCAGACCTGCACTCATGATGATGAGTCTGTGAGGGCGGAGATGATC 420
Oy 427 CGCTGGCAGCCAGCTGAGGAGCAGCCCGGCTCCCGGGCTCCGTAACCTGCTGATGTT 486
Db 421 CGCTGGCAGCCAGCTGAGGAGCAGCCCGGCTCCCGGGCTCCGTAACCTGCTGATGTT 480
Oy 487 TCTACACGAGAAAGAGAGTCTGAGCCAGAGATGAGACGGTCTCTGGGCGTGAATTT 546
Db 481 TCTACACGAGAAAGAGAGTCTGAGCCAGAGATGAGACGGTCTCTGGGCGTGAATTT 540
Oy 547 CCTGACAGAGCTCCCGCAGCTCAGCTGAGGCGTGTCTTGGCCCTCTGAGATGAGACC 606
Db 541 CCTGACAGAGCTCCCGCAGCTCAGCTGAGGCGTGTCTTGGCCCTCTGAGATGAGACC 600
Oy 607 CAGGTACTTAGATGAGAGCGGGGCTTACGCTGACGCTGCTGAGGCAAGACCGGATC 666
Db 601 CAGGTACTTAGATGAGAGCGGGGCTTACGCTGACGCTGCTGAGGCAAGACCGGATC 660

Oy 667 TTCAAGCCCATCTCCATCCAGACATGTGGGCCACATCTCCAGATATTGCACCAAGCATGT 726
Db 661 TTCAAGCCCATCTCCATCCAGACATGT----- 688
Oy 727 GAGGAGGCTTAGAGAGGCGGCTTTACCGGGGTGCAGTGCCCTTACCTGGGCGAGCCAC 786
Db 689 ----- 688
Oy 787 TACCAAGAGAGACTGAACTCCGACACAGAGCTGCCAATGATGATGAGCAGGCTATGGCCGAC 846
Db 689 ----- 688
Oy 847 CTGAGTCTTGGCGGCTTCCAGCGCCAGCCCTGGCGGGTCTTCAGAACAGAGAGCATG 906
Db 689 ----- -GGTCTGAGAACAGAGAGCATG 711
Oy 907 GAGCAGGCGATCGTGTGAGCTGTGAAAGTGTGATGTAGTACCTGAGAGTGTG 966
Db 712 GAGCAGGCGATCGTGTGAGCTGTGAAAGTGTGATGTAGTACCTGAGAGTGTG 771
Oy 967 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC 1026
Db 772 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC 831
Oy 1027 CGTGAATTCATGACACACAGATGCTGCTGTGGTGGGACACAGGGAGCCGAGCCTCCCG 1086
Db 832 CGTGAATTCATGACACACAGATGCTGCTGTGGTGGGACACAGGGAGCCGAGCCTCCCG 891
Oy 1087 ATCTTCCCGCCACCTTACCTGAGGCTCAGAGTGAACGCGACGAACCTGTGAGAGAGCTGAC 1146
Db 892 ATCTTCCCGCCACCTTACCTGAGGCTCAGAGTGAACGCGACGAACCTGTGAGAGAGCTGAC 951
Oy 1147 AGGAACAGGCTACCCACATCTTGAACATGSCCCGGAGATTGACAACTTCTACCTGAG 1206
Db 952 AGGAACAGGCTACCCACATCTTGAACATGSCCCGGAGATTGACAACTTCTACCTGAG 1011
Oy 1207 CGCTTACCTTACACAAATGTGGGCTGTGGATGAGGATGAGGATGAGGATGAGTGTGCGGAC 1266
Db 1012 CGCTTACCTTACACAAATGTGGGCTGTGGATGAGGATGAGGATGAGGATGAGTGTGCGGAC 1071
Oy 1267 TGGAAAGAGAGCAGCGCTTATGAGGCTGCAAGAGCAGAGGACCCACAGCTGTGATC 1326
Db 1072 TGGAAAGAGAGCAGCGCTTATGAGGCTGCAAGAGCAGAGGACCCACAGCTGTGATC 1131
Oy 1327 CACTGCAAGATGGGCGTCAAGCGCTCAGCGGCACAGTCTGTGCTTATGCAATGAAGCAG 1386
Db 1132 CACTGCAAGATGGGCGTCAAGCGCTCAGCGGCACAGTCTGTGCTTATGCAATGAAGCAG 1191
Oy 1387 TACGAATGAGCCTGAGAGAGGCGGCGGCGCAGGTCAGAGAGTCCGGCCATCGGCCGC 1446
Db 1192 TACGAATGAGCCTGAGAGAGGCGGCGGCGCAGGTCAGAGAGTCCGGCCATCGGCCGC 1251
Oy 1447 CCCAACCTGGGCTTCTTGGCCAGCTGCAAGATCTTACAGAGGCAATCTTGAAGGCGCAAC 1506
Db 1252 CCCAACCTGGGCTTCTTGGCCAGCTGCAAGATCTTACAGAGGCAATCTTGAAGGCGCAAC 1311
Oy 1507 TGAAGGTGTGGGAGAGAGAGTGTAGGCAATGGAAGAGAGCAGGAGCCCGCAAA 1566
Db 1312 TGAAGGTGTGGGAGAGAGAGTGTAGGCAATGGAAGAGAGCAGGAGCCCGCAAA 1371
Oy 1567 AGAGCCTGGG - -CCAGGGCCAGGTATAACCTCCGAGGGGTCTAGAGGTCATCACTT 1624
Db 1372 AGAGCCTGGGCGCAGGGGCGCTATTAACCTCCGAGGGGTCTAGAGGTCATCACTT 1431
Oy 1625 CTGAGCCCTCTCTT -GGAGCTGAGAGCAGCTCAG -AGACCACTGACATGCCAGAGTCT 1682
Db 1432 CTGAGCCCTCTCTTGGGAGAGTGGAGAGCACTCAGTAGAACCACTGACATGCCAGAGTCT 1491
Oy 1683 TCTCTTCCAGAGAGTCTTCAATGAAGAGCTTGCAGGCCCTTCCACAGCTTGCAGAG 1742
Db 1492 TCTCTTCCAGAGAGTCTTCAATGAAGAGAGCTTGCAGGCCCTTCCACAGCTTGCAGAG 1551
Oy 1743 CCAAGGAGCGCAGCAGGTGACAGGGGCGCTCAGCCTGCCGTAAGTCCCGCAGTCA 1802

934 AAGTGTGATGTCAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACGGCTCTG 993
| | | | |
841 AAGTGTGATGTCAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACGGCTCTG 900
| | | | |
994 GAGCTGCGCTGGGGCTCCCTCCAGAGTACCGTGAATTCATGCACCAACAGATGCTG 1053
| | | | |
901 GAGCTGCGCTGGGGCTCCCTCCAGAGTACCGTGAATTCATGCACCAACAGATGCTG 960
| | | | |
1054 CTGCTGTGGCAGCAGGGAGCGAGCTCCCGATCTCCCGACCTCTACCTGGGGCTCA 1113
| | | | |
961 CTGCTGTGGCAGCAGGGAGCGAGCTCCCGATCTCCCGACCTCTACCTGGGGCTCA 1020
| | | | |
1114 GAGTGAAGCAGCAAACTGGAGAGCTGCAGAGAGAACAGGGGTACCCACATCTTTGAG 1173
| | | | |
1021 GAGTGAAGCAGCAAACTGGAGAGCTGCAGAGAGAACAGGGGTACCCACATCTTTGAG 1080
| | | | |
1174 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCGCTC 1233
| | | | |
1081 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCGCTC 1140
| | | | |
1234 TGGGATGAGAGTGCAGCGCAGCTGCTGCCGCACTGGAGAGAGACCGCTTCATTTAG 1293
| | | | |
1141 TGGGATGAGAGTGCAGCGCAGCTGCTGCCGCACTGGAGAGAGACCGCTTCATTTAG 1200
| | | | |
1294 GCTGCAAGAGCAGAGGGACCCACAGTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1353
| | | | |
1201 GCTGCAAGAGCAGAGGGACCCACAGTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1260
| | | | |
1354 GCGGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
| | | | |
1261 GCGGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
| | | | |
1414 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
| | | | |
1321 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
| | | | |
1474 CAGATCTACAGGGCATCTCTGACG----- 1497
| | | | |
1381 CAGATCTACAGGGCATCTCTGACG----- 1440
| | | | |
1498 ----- 1497
| | | | |
1441 GGTGGGCTCTCCCGAGAGAGCAGCCGCCCTGAAGTCTTACACCATTTCCACCTCTT 1500
| | | | |
1498 - - -GCCAGAACTGAGGGTGGTGGGGAGAGAACTTGTAGCATGGAAGAGAGCCAGCA 1555
| | | | |
1501 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
| | | | |
1556 GCCCGAAGAGAGAGCTGCGGCGCAGCGCAGCTATATAACCTCCGAGGGGTCAATGAGGTCC 1615
| | | | |
1561 GCCCGAAGAGAGAGCTGCGGCGCAGCGCAGCTATATAACCTCCGAGGGGTCAATGAGGTCC 1620
| | | | |
1616 ATCAGCTCTTGAAGCGCTCTGAGAGCTGAGAGAGCACTCAGAGAGCAAGTACATGCCA 1675
| | | | |
1621 ATCAGCTCTTGAAGCGCTCTGAGAGCTGAGAGAGCACTCAGAGAGCAAGTACATGCCA 1680
| | | | |
1676 GAGGTCTTCTCTCCACAGAGTCTTACATGAAGAGAGCTCTGAGAGCTTCCACAGCTT 1735
| | | | |
1681 GAGGTCTTCTCTCCACAGAGTCTTACATGAAGAGAGCTCTGAGAGCTTCCACAGCTT 1740
| | | | |
1736 GCAAG 1795
| | | | |
1741 GCAAG 1800
| | | | |
1796 CAGTCAAGTGTACCTCCAG 1855
| | | | |
1801 CAGTCAAGTGTACCTCCAG 1860
| | | | |
1856 GAGCAG 1915
| | | | |
1861 GAGCAG 1920
| | | | |

994 AAGTGTGATGTCAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACGGCTCTG 993
| | | | |
841 AAGTGTGATGTCAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACGGCTCTG 900
| | | | |
994 GAGCTGCGCTGGGGCTCCCTCCAGAGTACCGTGAATTCATGCACCAACAGATGCTG 1053
| | | | |
901 GAGCTGCGCTGGGGCTCCCTCCAGAGTACCGTGAATTCATGCACCAACAGATGCTG 960
| | | | |
1054 CTGCTGTGGCAGCAGGGAGCGAGCTCCCGATCTCCCGACCTCTACCTGGGGCTCA 1113
| | | | |
961 CTGCTGTGGCAGCAGGGAGCGAGCTCCCGATCTCCCGACCTCTACCTGGGGCTCA 1020
| | | | |
1114 GAGTGAAGCAGCAAACTGGAGAGCTGCAGAGAGAACAGGGGTACCCACATCTTTGAG 1173
| | | | |
1021 GAGTGAAGCAGCAAACTGGAGAGCTGCAGAGAGAACAGGGGTACCCACATCTTTGAG 1080
| | | | |
1174 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCGCTC 1233
| | | | |
1081 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCGCTC 1140
| | | | |
1234 TGGGATGAGAGTGCAGCGCAGCTGCTGCCGCACTGGAGAGAGACCGCTTCATTTAG 1293
| | | | |
1141 TGGGATGAGAGTGCAGCGCAGCTGCTGCCGCACTGGAGAGAGACCGCTTCATTTAG 1200
| | | | |
1294 GCTGCAAGAGCAGAGGGACCCACAGTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1353
| | | | |
1201 GCTGCAAGAGCAGAGGGACCCACAGTGTGCTGCTCAGAGATGGGCGCTCAGCGCTCA 1260
| | | | |
1354 GCGGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
| | | | |
1261 GCGGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
| | | | |
1414 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
| | | | |
1321 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
| | | | |
1474 CAGATCTACAGGGCATCTCTGACG----- 1497
| | | | |
1381 CAGATCTACAGGGCATCTCTGACG----- 1440
| | | | |
1498 ----- 1497
| | | | |
1441 GGTGGGCTCTCCCGAGAGAGCAGCCGCCCTGAAGTCTTACACCATTTCCACCTCTT 1500
| | | | |
1498 - - -GCCAGAACTGAGGGTGGTGGGGAGAGAACTTGTAGCATGGAAGAGAGCCAGCA 1555
| | | | |
1501 CCGCAGAGTGCAGAGTGCAGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
| | | | |
1556 GCCCGAAGAGAGAGCTGCGGCGCAGCGCAGCTATATAACCTCCGAGGGGTCAATGAGGTCC 1615
| | | | |
1561 GCCCGAAGAGAGAGCTGCGGCGCAGCGCAGCTATATAACCTCCGAGGGGTCAATGAGGTCC 1620
| | | | |
1616 ATCAGCTCTTGAAGCGCTCTGAGAGCTGAGAGAGCACTCAGAGAGCAAGTACATGCCA 1675
| | | | |
1621 ATCAGCTCTTGAAGCGCTCTGAGAGCTGAGAGAGCACTCAGAGAGCAAGTACATGCCA 1680
| | | | |
1676 GAGGTCTTCTCTCCACAGAGTCTTACATGAAGAGAGCTCTGAGAGCTTCCACAGCTT 1735
| | | | |
1681 GAGGTCTTCTCTCCACAGAGTCTTACATGAAGAGAGCTCTGAGAGCTTCCACAGCTT 1740
| | | | |
1736 GCAAG 1795
| | | | |
1741 GCAAG 1800
| | | | |
1796 CAGTCAAGTGTACCTCCAG 1855
| | | | |
1801 CAGTCAAGTGTACCTCCAG 1860
| | | | |
1856 GAGCAG 1915
| | | | |
1861 GAGCAG 1920
| | | | |

RESULT 6
US-09-955-732-20
; Sequence 20, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-955-732-20

Query Match 52.48; Score 1416; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
94 ATGGCCCTGTGTCAGAGTGAAGCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 153
| | | | |
1 ATGGCCCTGTGTCAGAGTGAAGCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
| | | | |
154 CCCTGGAGACAGGCGGCTCCAGGAGAGAGTGCAGTCCAGGAGAGAGAGAGAGAGAGAGAGAG 213
| | | | |
61 CCCTGGAGACAGGCGGCTCCAGGAGAGAGTGCAGTCCAGGAGAGAGAGAGAGAGAGAGAGAG 120
| | | | |
214 CTCGCGGGGCTGCTCTGAGAGTGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 273
| | | | |
121 CTCGCGGGGCTGCTCTGAGAGTGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
| | | | |
274 AGTTCTGAGACCAAG 333
| | | | |
181 AGTTCTGAGACCAAG 240
| | | | |
334 TTCCGGGCAAGAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
| | | | |
241 TTCCGGGCAAGAGATCCCGAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
| | | | |
394 ATGTTACAGCTGCTGAGAGCGGAGAGATGATCCGCTGCGAGAGAGAGAGAGAGAGAGAGAG 453
| | | | |
301 ATGTTACAGCTGCTGAGAGCGGAGAGATGATCCGCTGCGAGAGAGAGAGAGAGAGAGAGAG 360
| | | | |
454 GGGCGTCCCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
| | | | |
361 GGGCGTCCCGGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
| | | | |
514 CAGATGAGAGAGAGTCCCTGAGAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
| | | | |
421 CAGATGAGAGAGAGTCCCTGAGAGTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
| | | | |
574 CTGGGCTGTGCTTCCCTGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
| | | | |
481 CTGGGCTGTGCTTCCCTGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
| | | | |
634 TTACAGGTGAGCTGTGAG 693
| | | | |
541 TTACAGGTGAGCTGTGAG 600
| | | | |
694 TTGGGCGACACTCAGAGTATTCAG 753
| | | | |
601 TTGGGCGACACTCAGAGTATTCAG 660
| | | | |
754 CCGGCTGTGAGTCCCTGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 813
| | | | |


```

|||||
Db 661 CCGGATGGACAGCCCTCACTGGGCGCAGCCATACAGAGAGAGCTAAGTCCGACAG 720
Qy 814 AGCTGCCATATAGTGGAGGCTATGCGGACCTGGAGTCTCTGGCCCTCCAGGCC 873
Db 721 AGCTGCCATATAGTGGAGGCTATGCGGACCTGGAGTCTCTGGCCCTCCAGGCC 780
Qy 874 GAGCTGGCGGCTCTCTCAAGACAGAGAGAGAGAGGAGTCCGCTGAGTGTGG 933
Db 781 GAGCTGGCGGCTCTCTCAAGACAGAGAGAGAGAGGAGTCCGCTGAGTGTGG 840
Qy 934 AAGTGTGATGTCAGTACCTGGAGAGTCTCACTTCAAGAGATCCGCGAGCTTG 993
Db 841 AAGTGTGATGTCAGTACCTGGAGAGTCTCACTTCAAGAGATCCGCGAGCTTG 900
Qy 994 GAGCTGGCGGCTCTCTCAAGACAGAGAGAGAGAGGAGTCCGCTGAGTGTGG 1053
Db 901 GAGCTGGCGGCTCTCTCAAGACAGAGAGAGAGAGGAGTCCGCTGAGTGTGG 960
Qy 1054 CTGCTGGTGGCAGACGCGGACCGAGACCTCCGCAATCTTCCGCACTTACCTGGAGTCA 1113
Db 961 CTGCTGGTGGCAGACGCGGACCGAGACCTCCGCAATCTTCCGCACTTACCTGGAGTCA 1020
Qy 1114 GAGTGAACGACAGAAACCTGGAGAGCTGCAGAGAAACAGGGTCCACCACATCTTGAC 1173
Db 1021 GAGTGAACGACAGAAACCTGGAGAGCTGCAGAGAAACAGGGTCCACCACATCTTGAC 1080
Qy 1174 ATGGCCCCGGGAGATTGACAACTTCTACCTGAGGCTTACCTACCAAAATGTGGCTTC 1233
Db 1081 ATGGCCCCGGGAGATTGACAACTTCTACCTGAGGCTTACCTACCAAAATGTGGCTTC 1140
Qy 1234 TGGATGAGAGTGGGCGCCAGCTGCTGCGCACTGGAAGAGAGACCGGCTTCAATGTAG 1293
Db 1141 TGGATGAGAGTGGGCGCCAGCTGCTGCGCACTGGAAGAGAGACCGGCTTCAATGTAG 1200
Qy 1294 GCTGCAAGAGACAGAGGACCCAGCTGCTGCTGCTCACTGCAAGATGGGCGTCAAGCCGTCA 1353
Db 1201 GCTGCAAGAGACAGAGGACCCAGCTGCTGCTGCTCACTGCAAGATGGGCGTCAAGCCGTCA 1260
Qy 1354 GCGGCGACAGTGGGCGCTATGCGCATGGAAGAGATGAGATGAGCTGGAGAGGCGCTG 1413
Db 1261 GCGGCGACAGTGGGCGCTATGCGCATGGAAGAGATGAGATGAGCTGGAGAGGCGCTG 1320
Qy 1414 CGCCACGTGACAGAGCTCCGCGCCATGCGCCGACCAACCTGAGCTTCTGCGCAGCTG 1473
Db 1321 CGCCACGTGACAGAGCTCCGCGCCATGCGCCGACCAACCTGAGCTTCTGCGCAGCTG 1380
Qy 1474 CAGATCTACCAAGGCACTCTGACGCGCAAGAACTGA 1509
Db 1381 CAGATCTACCAAGGCACTCTGACGCGCAAGAACTGA 1416

RESULT 7
US-09-775-925-1
: Sequence 1, Application US/09775925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Lucbe, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
: FILE REFERENCE: 200125, 420
: CURRENT APPLICATION NUMBER: US/09/775, 925
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1949
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-1

```

Query Match 12.5%, Score 336.8; DB 10; Length 1949;

Best Local Similarity 59.5%; Pred. No. 2,3e-81;
Matches 591; Conservative 0; Mismatches 397; Indels 6; Gaps 1;

```

Qy 517 GATGAGAGGCTCTCTGCGGCTGAGATTCCCTGACAGACAGCTCCCGACGTGACCTG 576
Db 32 GAGGAGATATCTGCTGCGGAGTGAATTTTCCAAATAGAAAGTAAAGTAAAGTAAAGTAA 91
Qy 577 GAGGAGATATCTGCTGCGGAGTGAATTTTCCAAATAGAAAGTAAAGTAAAGTAAAGTAA 636
Db 92 GAGGAGATATCTGCTGCGGAGTGAATTTTCCAAATAGAAAGTAAAGTAAAGTAAAGTAA 151
Qy 637 AGCGTACGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 696
Db 152 AGCGTACGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 211
Qy 697 GCGACACTCCAGCTATTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
Db 212 TCTGCCCTGACAGGCTCTTCAAGAGGCTGCGAAGTGGCCGCGAGGACACACTACTTCCC 271
Qy 757 GGTGAGTGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 816
Db 272 GGTGAGTGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 331
Qy 817 TGCTCAATGATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 876
Db 332 TGCTCAATGATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 391
Qy 877 C-----CTGGGAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 930
Db 392 CTATTTTGGAGACAAAGCCCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 451
Qy 931 TGGAAAGTGGTGGATGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 990
Db 452 TGGAAAGTGGTGGATGATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 511
Qy 991 CTGAGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1050
Db 512 TTAGAGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
Qy 1051 CTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1110
Db 572 CTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 631
Qy 1111 TCAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1170
Db 632 TCTGATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 691
Qy 1171 AACATGGCCCGGAGATGAGCAACTTACCTGAGCGCTTACCTGAGCAAGTGGCGC 1230
Db 692 AACATGGCCCGGAGATGAGCAACTTACCTGAGCGCTTACCTGAGCAAGTGGCGC 751
Qy 1231 CTCTGGGATGAGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1290
Db 752 GCTTACGATGAGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 811
Qy 1291 GAGGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1350
Db 812 AACAAAGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 871
Qy 1351 TCAGGCGGACAGTGGCTGCGCTATGCGCATGGAAGAGTGAAGTGAAGTGAAGTGAAG 1410
Db 872 TCGGCTCCACAGTCAATGAGCTATGCAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 931
Qy 1411 CTGGCGGACGTGACAGGAGTCCGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCG 1470
Db 932 TATAACTATGTAAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 991
Qy 1471 CTGCAATCTACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1504
Db 992 CTGCTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1025

```

RESULT 8

US-09-775-925-5
: Sequence 5, Application US/09775925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Lucche, Ralf M.
: APPLICANT: Mei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DNAL-SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASES
: FILE REFERENCE: 200125.420
: CURRENT APPLICATION NUMBER: US/09/775.925
: CURRENT FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1711
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-5

Query Match 11.6%; Score 313.4; DB 10; Length 1711;
Best Local Similarity 55.4%; Pred. No. 4.8e-75;
Matches 655; Conservative 0; Mismatches 516; Indels 12; Gaps 2;

QY 379 CAGCAGCTGCACCTGATGATGAGTGGCGGAGATGACATCGGCTGACAGCC 438
DB 348 CAGCTGTCCAGCAATGCTTCTTCTCTCCGCCAGAGACACATCAGGCTGCTGA 407
QY 439 CAGCTGAGGAGCCCGGCTCCCGGCTCCGCTACTCTGCTGATGTTCTACAGAGA 498
DB 408 AGACTGGAAAGTACTTACAGAAATGGAACAGCCTATATGATGTTTCACTAATGGT 467
QY 499 GGAGAGGCTGAGGCGAGAGTGAAGCGTCTCTGCGGCTGATTTCCCTGACAGCAG 558
DB 468 AGACAAG-----ACACTGAAGAAAGCATCGTCTAGGATGATTTCTCTATGAC 521
QY 559 TCCCCAGCTGCACCTGGGCTGCTGCTGCTGCTGAGAGACACCCAGGTGACTTA 618
DB 522 ACTGACACTGTGACATGAGGCTTATGTTGCTCTCTGAGGACACGCTAATTCATTGG 581
QY 619 GATGAGAGCGGGGCTGACGCTGACGCTGTGTTGGGCAAGCGGATCTTCAAGCCATC 678
DB 582 GATGCTGATGATGCTGCTGCTGCTGATGACAGGATACAGAGTACATATTAACCTGTA 641
QY 679 TCCATCCAGACCATGTTGGGCCACTCCAGTATTGACCAACAGATGTGAGCAGCTTA 738
DB 642 TCTGTGACAGCAATGTGGTCTGCTGACTACAGAGCTTACACAGGCTTGTGAAGTCCGACA 701
QY 739 GGCAGCGGCTGTACCGGCTGGAGTGGCCCTACCTGGGCGACGACATACAGAGAGA 798
DB 702 GGCATTAATCTACTACCCAGGACGCTATTCTCCTGCTGGGTGATTTATGAGAGCAT 761
QY 799 CTGAATCCGAGACAGATGCTCAATGATGAGAGGCTATGAGCGACCTGAGTCTCTG 858
DB 762 ATCACTGATGATCATCTCCATCATGATGATGATGATGATGATGATGATGATGATGAT 821
QY 859 CGGC-----CTCCAGGCGCGAGCGCTGGCGGCTCTCAAGAACAGAGACATGAGCAG 912
DB 822 CGGCGGAGCTCTCCAGCTCTCTTCAACCGACATACCTACTGACGTAAGCAAGCAAGAG 881
QY 913 GGGATCCGCTGAGGCTGGAAGTGTGATGTCAGTGAAGCTGAGAGTGTGACTTCC 972
DB 882 CTAAATTAACCAATTAAGGAGATGATGACAGAGAGATTTGAGATATTTATATCC 941
QY 973 AAAGAGATCCGCGAGGCTTGGAGCTGCGCTGGGCTCCCTCCAGAGATACCTGAC 1032
DB 942 AAAGAGATTAAGACAGAGTGTGAATGCAATGATGATGATGATGATGATGATGATGAT 1001
QY 1033 TTTATGACAAACAGATGCTGCTGTGTGTGAGCAGAGCGGAGCCGAGCTCCGATCTTC 1092
DB 1002 TTTATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1061
QY 1093 CCCACTCTACTGCTGGCTGACAGTGAAGCAGCAAACTGTGAGAGAGCTGACAGAGAAC 1152

DB 1062 GAGCATGTGTTCCGCGGCTCAGATGGAATGCTCTCAACTTAGAGAGATTACAGAACGA 1121
QY 1153 AGGCTACCCACATCTTGACATGATGCGCGGAGATTGACAACTTCTACCTGACGCTTC 1212
DB 1122 GGGTACGGTATATCTTGAATGTAATGAGATAGATAGTAACTTCTTCCAGAGCTTT 1181
QY 1213 ACCTAACCAATGTCGGCTGCTGGATGAGATGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1272
DB 1182 GAGTATCATTAATTCATTCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 1273 GAGAGCAGCCCTTCATTCATGAGCTGAGCAGAGACAGAGGACCCAGCTGCTGCTGCTGCT 1332
DB 1242 GACACTTACAAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1301
QY 1333 AAGATGGCGCTCAGCGCTGACGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1392
DB 1302 AAAATGGGGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
QY 1393 TGCAGCTGAGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
DB 1362 TGGATTCGAGACCGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1453 CTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1512
DB 1422 CCAAGCTTCATGAGAGCACTGGAAGATATCAGGGGATCTTGTGCTGCTGCTGCTGCTGCTG 1481
QY 1513 TGTGCGGAGAGAGGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1555
DB 1482 TTGATTCATGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1524

RESULT 9
US-09-969-708-398/c
: Sequence 398, Application US/09969708
: Patent No. US20020102532A1
: GENERAL INFORMATION:
: APPLICANT: Augustus, Meena
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-70
: CURRENT FILING DATE: 2001-10-03
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: US/60/237,606
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: US/60/237,425
: PRIOR FILING DATE: 2000-10-03
: NUMBER OF SEQ ID NOS: 658
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 398
: LENGTH: 334
: TYPE: DNA
: ORGANISM: Homosapiens
US-09-969-708-398

Query Match 11.6%; Score 313.2; DB 10; Length 334;
Best Local Similarity 96.1%; Pred. No. 2.8e-75;
Matches 321; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2370 TGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429
DB 334 TGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
QY 2430 CCCCAAGGAGGAGATCTCAGAGCCCGGAGGAGGCTGGAAGGCTGAGAT 2489
DB 274 CCCCAAGGAGGAGATCTCAGAGCCCGGAGGAGGCTGGAAGGCTGAGAT 215
QY 2490 GCGTTCCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2549
DB 214 CGTTCCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 155
QY 2550 ACACCAAGCAGATCATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2609

US-09-865-993-10

Query Match	3.3%	Score 89.2;	DB 9;	Length 2479;
Best Local Similarity	51.5%;	Pred. NO. 4e-14;		
Matches 205; Conservative	0;	Mismatches 193;	Indels 0;	Gaps 0;

OY	1087	ATCTTCCCCACGCTTACCTGAGGCTCAGAGTGGAAAGCAGCAACCTGGAGAGCTGCAG	1146
Db	758	ATCTCTCCCTCCTCTTACTTGGAAAGTGCCTTACATGATCTCCAAAGTGGAGTCTCTGC	817
OY	1147	AGGAACAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGACAACTTCTACCCCTGAG	1206
Db	818	AACTTGCACATCAGAGCCCTGCTGAAATGTCTCCGACGGAGACTCCGAGGCTCGATGAC	877
OY	1207	CGCTTACCTTACCAACATGTGCGCTCTCTGGATGAGGAGTGGCCAGCTGCTGCGCAC	1266
Db	878	CACCTACACTACAAATGGATCCCTGTGGAAAGACAGCAACCGCTGACATTAGCTCCAC	937
OY	1267	TGGAAGGAGAGCGACCGCTCATTTGAGGGCTGCAAGAGCAGCAGGCAACCACTGCTGTGCT	1326
Db	938	TTTTCAGAAAGCAATAGACTCTTATGTAGCTGTGTGAGGAAAAGGAGGCAAGTCTCTGTCT	997
OY	1327	CACGTCAAGATGGGCGTCAGCCGCTCAGCGGCACAGTGTCTGGCTATGCATTGAAGAG	1386
Db	998	CACGTGAGGCTTGGAGATCTCCGTTACACCCACCATGTGATGTGCTTACTTATGAAGAC	1057
OY	1387	TACGATGCAAGCTGGAGCAGCGCCCTGGGCGCAAGTGCAGGAGCTCCGGCCATCGCCGC	1446
Db	1058	AAGCAGTTCGCGCTGGAAGGAGGCGCTTCGATTATCAACACAGAGGAGCATGGTCTCG	1117
OY	1447	CCCAACCTGGCTTCTCGCGCCAGCTGCAGATCTACCA	1484
Db	1118	CCCACTTTGGCTTATAGGGCCAGCTCTCTGCAGATGAGA	1155

RESULT 15
US-09-902

```

Sequence 1918, Application US/09902941
Patent No. US20020172952A1
GENERAL INFORMATION:
APPLICANT: Henderson, Robert A.
APPLICANT: Wang, Tongtong
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Banjur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C17
CURRENT APPLICATION NUMBER: US/09/902,941
CURRENT FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 2002
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1918
LENGTH: 1209
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-941-1918

```

Query Match	3.38;	Score 88;	DB 9;	Length 1209;
Best Local Similarity	50.58;	Pred. No. 6.3e-14;		
Matches 214;	Conservative 0;	Mismatches 210;	Indels 0;	Gaps 0;

QY	1087	ATCTTCCCAACCTTACCTTACCTTGGGCTAGAGTGGAAACGACGACAAACCTGGAGAGCTGGAC	1146
Db	616	ATCTTCCCTTCTCTTACCTTCGGCACTGCCCTACCACTGCTGCTCCCGGAGMACATCTGCTGGAC	675
QY	1147	AGGAACAGGGTCACCCACATCTTGAACATATGGCCCGGAGATTTGACACATCTTACACCTTAC	1208

Db	676	GCCTGGGACACAGGCTCTGTGAATGTCCTCTCGAATGCCCCAACACACTTTGAAGCA	735
QY	1207	GCCTTACCTACCAACATGTGCGCCCTCTGGATGAGAGTGGCCAGCTGCTGCCGAC	1266
Db	736	CACATACAGTACAAGTACATCCAGTGAAGATAACCAACAAAGCCGACATCAGCTCTGG	795
QY	1267	TGGAAGAGACGCAACCCCTCTATTGAGGCTGCAAGAGCACAGGCAACCCAGTGTGTC	1326
Db	796	TTCTATGAGAGCCATAGAGTACATGATGCCGTGAAGAGCTGCCGGGGGCGCTGTGTG	855
QY	1327	CACGTACAAGATGGGCGTCACGCCCTCAGCGGCGCACAGTGTCTGGCTATGCCATGAAGAG	1386
Db	856	CACGTCCAGGCGGGCATTCTCGCGGTGGGCGCACATCTGCTGCTTACCTGATGTAAAG	915
QY	1387	TACGATGCAAGCTTGGAGCAGGGCCCTGCGCCACAGTCCACAGAGACTTCGGGCCATCGCCCC	1446
Db	916	AAACGGGTGAGGGCTGGAGGAGGGCCCTTCGATTCGTTAAACACAGCGCGCAGCATCTATCG	975
QY	1447	CCCAACCTTGCTTCGCGCCACAGCTAGCATATACACAGGCGATCCTGACGCGCAAGAC	1506
Db	976	CCCAACTTCAGCTTCAGTGGGCGAGCTGCTGCAATTCGATTCACAGAGTGTGTCACAGTCC	1035
QY	1507	TGAG 1510	
Db	1036	TGTG 1039	

Search completed: April 12, 2003, 07:29:43
Job time : 426.978 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:58:55 ; Search time 7346.79 Seconds

(without alignments)
5960.776 Million cell updates/sec

Title: US-09-761-640-1

Perfect score: 2704
Sequence: 1 cgtccctctgctcgtcg...aaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	839	31.0	991	13	BM471207 AGENCOURT
2	767.2	28.4	875	12	BG752969 602732427
3	758	28.0	812	12	BG750408 602709243
4	755.2	27.9	790	12	BG326709 602425578
5	741.6	27.4	772	14	BQ014446 UT-H-ED1-
6	727.8	26.9	1131	12	BE737389 601306112

7	698.6	25.8	846	10	BE563259
8	695.6	25.7	932	12	BE340220
9	673.8	24.9	726	12	BE869222
10	664	24.6	910	12	BG169297
11	662	24.5	663	12	BG753072
12	652.4	24.1	1004	14	BM910493
13	642	23.7	653	12	BG679981
14	641.6	23.5	719	12	BE737087
15	635.4	23.3	716	12	BE907514
16	622	23.0	890	12	BE901875
17	619.2	22.9	897	12	BE873337
18	609.4	22.5	672	12	BG831015
19	600.4	22.2	1228	14	BM921421
20	599.4	22.2	1125	14	BM809903
21	599	22.2	598	10	BE280203
22	598.4	22.1	678	13	BM047611
23	596	22.0	1024	13	BM560192
24	590.4	21.8	622	12	BE898514
25	586.4	21.7	589	9	AL602806
26	577.6	21.4	1061	14	BM906413
27	567.2	21.0	728	13	BI824278
28	562.6	20.8	791	12	BE881145
29	561.8	20.8	892	12	BE971565
30	548.8	20.3	917	14	BO712114
31	548.4	20.3	584	13	BM018531
32	541.6	20.0	863	12	BG749236
33	537.6	19.9	969	10	BE563116
34	535.4	19.8	943	13	BM563401
35	517.4	19.1	574	13	BI007055
36	517	19.1	937	14	BO718968
37	516	19.1	719	10	BE543087
38	505	18.7	750	10	BE531347
39	502	18.6	846	13	BI820326
40	490.2	18.1	560	13	BI007133
41	482.4	17.8	568	13	BI087083
42	481.6	17.8	496	10	BE348287
43	471	17.4	505	13	BM083928
44	469.8	17.4	952	14	BO675874
45	468	17.3	729	13	BI086720

ALIGNMENTS

RESULT 1
BM471207
LOCUS
DEFINITION
AGENCOURT.6478155 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:5563111
5', mRNA sequence.
BM471207
BM471207.1 GI:18520249
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 991)
NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LMNL293 row: j column: 08
High quality sequence stop: 592.
Location/Qualifiers
1. .991

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:556311"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/site="Organ: skin; Vector: pCMV-Sport6; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

```

```

BASE COUNT      199 a      303 c      329 g      159 t      1 others
ORIGIN

```

```

Query Match      31.0%; Score 839; DB 13; Length 991;
Best Local Similarity 93.5%; Pred. No. 1,1e-116;
Matches 920; Conservative 0; Mismatches 56; Indels 8; Gaps 4;

```

```

QY 40 GTTAGAGGAAGGGCCCTGGCCCGTGCACCCAGAGTGTCTGGGCTGCTCCATGGCC 99
DB 1 GTTAGAGGAAGGGCCCTGGCCCGTGCACCCAGAGTGTCTGGGCTGCTCCATGGCC 60
QY 100 CTGTGTCACATGAGCCCTTGGCCCGGCGAGCGGCTCTCCAGCCCGTGGGGCCCTG 159
DB 61 CTGTGTCACATGAGCCCTTGGCCCGGCGAGCGGCTCTCCAGCCCGTGGGGCCCTG 120
QY 160 GACCAAGCGGTCCAGCAAGAGTCCAGTCCAGCAAGAGAGCTTTGGGTGCTCCGT 219
DB 121 GACCAAGCGGTCCAGCAAGAGTCCAGTCCAGCAAGAGAGCTTTGGGTGCTCCGT 180
QY 220 GGGGCTGTCTGGAGCTGACGATGAGGAGGAGCAATGATGATGATGATGATGATG 279
DB 181 GGGGCTGTCTGGAGCTGACGATGAGGAGGAGCAATGATGATGATGATGATGATG 240
QY 280 GAGCCAAAGAGAAGGCCCGGAGTGGAGAGAGCTCCAGGGGACAGACAGATTGCG 339
DB 241 GAGCCAAAGAGAAGGCCCGGAGTGGAGAGAGCTCCAGGGGACAGACAGATTGCG 300
QY 340 CAAGGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
DB 301 CAAGGATCCCAAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 400 CAGCTGCTGAGGCGCGAGATGACATCCCGCTGGCAAGCCAGCTGAGAGCAGCCG 459
DB 361 CAGCTGCTGAGGCGCGAGATGACATCCCGCTGGCAAGCCAGCTGAGAGCAGCCG 420
QY 460 CCGCGGCTCCGCTACCTGCTGAGTCTTCTACAGAGAGAGAGAGAGAGAGAGAG 519
DB 421 CCGCGGCTCCGCTACCTGCTGAGTCTTCTACAGAGAGAGAGAGAGAGAGAGAG 480
QY 520 GAGAGGAGTCTGAGGAGTGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAG 579
DB 481 GAGAGGAGTCTGAGGAGTGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAG 540
QY 580 CTGTGCTTGGCCCTCTGAGAGTGAACACCGAGTGTACTAGATGAGAGAGAGAG 639
DB 541 CTGTGCTTGGCCCTCTGAGAGTGAACACCGAGTGTACTAGATGAGAGAGAGAG 600
QY 640 GTGACGCTGTGGTGGGCAAGCCGAGATCTTCAAGCCATCTCCATCCAGACATG 699
DB 601 GTGACGCTGTGGTGGGCAAGCCGAGATCTTCAAGCCATCTCCATCCAGACATG 660
QY 700 ACATCCAGATTTGACACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
DB 661 ACATCCAGATTTGACACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 759 TGGAGAGTCCCTACCTGGGCGACCACTACAGAGAGAGAGAGAGAGAGAGAGAG 818
DB 721 TGGAGAGTCCCTACCTGGGCGACCACTACAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 819 CTTCAATGAGTGGAGCGCTATGGCCGAGCTGTGAGTCTCTGGCTCCAGCGCCGAG 878
DB 781 CTTCAATGAGTGGAGCGCTATGGCCGAGCTGTGAGTCTCTGGCTCCAGCGCCGAG 840

```

```

QY 879 TGGCGGCTCTCAG-AACAGAGACAGATGAGAGAGAGAGAGAGAGAGAGAGAG 934
DB 841 TGGCGGCTCTCCGAAACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 935 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 991
DB 901 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 992 TGGAGCTGAGCTGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015
DB 961 TGGAGCTGAGCTGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984

```

```

RESULT 2
BG752969
LOCUS
DEFINITION
602732427F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876014 5',
mRNA sequence.
ACCESSION
BG752969
VERSION
BG752969.1 GI:14063622
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 875)
NIH-MGC http://mgi.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM1758 row: e column: 07
High quality sequence stop: 818.
Location/Qualifiers
1. 875

```

```

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4876014"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(5). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1"

```

```

BASE COUNT      184 a      267 c      275 g      149 t
ORIGIN

```

```

Query Match      28.4%; Score 767.2; DB 12; Length 875;
Best Local Similarity 97.0%; Pred. No. 6.7e-106;
Matches 846; Conservative 0; Mismatches 18; Indels 8; Gaps 6;

```

```

QY 289 GAGAAAGCCCCGAGTAGAGAGAGTCCAGGGGACACAGAGAGTGGGCAAGAGATCC 348
DB 1 GAGAAAGCCCCGAGTAGAGAGAGTCCAGGGGACACAGAGAGTGGGCAAGAGATCC 60
QY 349 CAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
DB 61 CAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 409 AGCGCGAGAGATGACATCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468

```


|||||
Db 661 AGTGAAGTGGAGAGTGTACTTCCAAAGAGATCCGACAGGCTCTGGAGCTTGGCGCTGGG 720
OY 1009 CTCCTCCCTCCACAGTACGAGTGTGACTGATCGACACACCAATGCTGCTGGTGGACAG 1068
Db 721 TGTCCCTCCCTCCACAGTACGAGTGTGACTGATCGACACCAATGCTGCTGGTGGACAG 779
OY 1069 CGGAGCCGAGCTCCGCTGCTTCCCTCCCTGCT 1102
Db 780 C-GGAACGAGGCTCCCGATTTCCTCCACTCTACT 812

RESULT 4
BG326709 790 bp mRNA linear EST 27-FEB-2001
LOCUS 603425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
DEFINITION mRNA sequence.
ACCESSION BG326709
VERSION BG326709.1 GI:13133146
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 790)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10M1277 row: k column: 20
High quality sequence stop: 737.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4563595"
/clone_lib="NIH_MGC_14"
/lisse_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 162 a 237 c 268 g 123 t
ORIGIN

Query Match 27.9% Score 755.2: DB 12: Length 790;
Best Local Similarity 98.2%: Pred. No. 4.4e-104;
Matches 774: Conservative 0; Mismatches 13; Indels 1; Gaps 1;

OY 45 GGAAGAGGCGCGTGCCTCCGACCCAGTGTGCGGCTGCTCCATGAGCTGCT 104
Db 1 GGAAGAGGCGCGTGCCTCCGACCCAGTGTGCGGCTGCTCCATGAGCTGCT 60
OY 105 CACAGTGAAGCGTTGCCCCCGGAGCGGCGCTTCACAGCCCTGGGGCCCTGGACCA 164
Db 61 CACAGTGAAGCGTTGCCCCCGGAGCGGCGCTTCACAGCCCTGGGGCCCTGGACCA 120
OY 165 GCGGCTCAGGAGAGTGCAGTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 224
Db 121 GCGGCTCAGGAGAGTGCAGTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 180

OY 225 TGNCTGGGAGTGCAGATGAGAGAGGACAATGATGACAGCAGAGCCAGTTCTAGCC 284
Db 181 TGTCTGGGAGTGCAGATGAGAGAGGACAATGATGACAGGAGCCAGTTCTAGCC 240
OY 285 AACAGAGAGGCGCCGAGTGAAGAGAGAGTCCACGCGGAGCAGACAGACTTGGGCAAG 344
Db 241 AACAGAGAGGCGCCGAGTGAAGAGAGAGTCCACGCGGAGCAGACAGACTTGGGCAAG 300
OY 345 ATCCAGAGTCCCGAG 404
Db 301 ATCCAGAGTCCCGAG 360
OY 405 GCTGAGCGCCGAGAGATGATCCGCTGGCAGCCGAGCGGAGGAGCCCGCGCTCCCG 464
Db 361 GCTGAGCGCCGAGAGATGATCCGCTGGCAGCCGAGCGGAGGAGCCCGCGCTCCCG 420
OY 465 GCTCCGCTACCTGCTGCTGATGTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 524
Db 421 GCTCCGCTACCTGCTGCTGATGTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 480
OY 525 GGTCTCTCTGGGCGTGGATTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
Db 481 GGTCTCTCTGGGCGTGGATTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
OY 585 CTTCGCGCTGAG 644
Db 541 CTTCGCGCTGAG 600
OY 645 GTCTGTGGGCAAGCGGAGTCTTCAAGCCCATCTCCATCCAGAGAGAGAGAGAGAG 704
Db 601 GTCTGTGGGCAAGCGGAGTCTTCAAGCCCATCTCCATCCAGAGAGAGAGAGAGAG 660
OY 705 CCAGGATTTGACAC -AAGCATGTGAGGAGAGCTGTAGGAGCGGCTTGTACCGGCTGCA 763
Db 661 CCAGGATTTGACAC -AAGCATGTGAGGAGAGCTGTAGGAGCGGCTTGTACCGGCTGCA 720
OY 764 GTGCCCTCACTGGGCGCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 823
Db 721 GTAGCCTCACTGGGCGCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
OY 824 ATGAGTGG 831
Db 781 ATGAGTGG 788

RESULT 5
B0014446/c 772 bp mRNA linear EST 26-MAR-2002
LOCUS B0014446
DEFINITION UI-H-ED1-act-9-13-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone
IMAGE:5833380 3', mRNA sequence.
ACCESSION B0014446
VERSION B0014446.1 GI:19739347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 772)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
POLY-A=yes.

Db 125 CTCCTGACACTGAGAGATCAACAACCTCTTGAGAGAACCCCTCAGCTGTGGC 184
QY 2055 GCACACATCTCTCTGAGCTCCGCCATACCCGCTACTACAGCCTCAGCTCCACCCCTG 2114
Db 185 GCACACATCTCTCTGAGCTCCGCCATACCCGCTACTACAGCCTCAGCTCCACCCCTG 244
QY 2115 TCACACTGAGGCTCAGCTCCCAACCCCTGTACTACAGCCTCAGCTCCACCCCTAAGTC 2174
Db 245 TCACACTGAGGCTCAGCTCCCAACCCCTGTACTACAGCCTCAGCTCCACCCCTAAGTC 304
QY 2175 CCAGGCCATCTCTGCTCCAGGCTCAAGACTTTCTAAGTGTAGAGAGG 2234
Db 305 CCAGGCCATCTCTGCTCCAGGCTCAAGACTTTCTAAGTGTAGAGAGG 364
QY 2235 ACTGAGGTACTCTTGGGGGACAGACACCCAGTATTCTCAACTCTAGCCTGAC 2294
Db 365 ACTGAGGTACTCTTGGGGGACAGACACCCAGTATTCTCAACTCTAGCCTGAC 424
QY 2295 ACTCAGCTGTGGCAGGATGAAAACAGAGCTTCCGTCAGAAAAGGTCACGCTCCCA 2354
Db 425 ACTCAGCTGTGGCAGGATGAAAACAGAGCTTCCGTCAGAAAAGGTCACGCTCCCA 484
QY 2355 CCCCCCCCCCTCCCTGACACCTCTGCTCCAGTTCATCTCGAAGCAGCCAGC 2414
Db 485 CCCCCCCCCCTCCCTGACACCTCTGCTCCAGTTCATCTCGAAGCAGCCAGC 544
QY 2415 CAGGCAACCAATGAGCCCAAGAGGAGAGATCTCAGGCCCCAGCCGAGAGGCTG 2474
Db 545 CAGGCAACCAATGAGCCCAAGAGGAGAGATCTCAGGCCCCAGCCGAGAGGCTG 604
QY 2475 CAGGAGCTGAGAGATGCTCTCTCTATCAGCTCAGGCTCAGAGTCTTCTCTGTC 2534
Db 605 CAGGAGCTGAGAGATGCTCTCTCTATCAGCTCAGGCTCAGAGTCTTCTCTGTC 663
QY 2535 CCCAGACCTCTCTGAGCAGCAGCCAGATCACAGGCGCAGAGATGATGCTCTCT 2594
Db 664 CCCAGACCTCTCTGAGCAGCAGCCAGATCACAGGCGCAGAGATGATGCTCTCT 721
QY 2595 TTTGCTCTTCTGAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2654
Db 722 TTTGCTCTTCTGAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 781
QY 2655 ACTGGAATTAACATTTTCAATTAACATTAACATTAACATTAACATTAACATTAAC 2697
Db 782 ACTGGAATTAACATTTTCAATTAACATTAACATTAACATTAACATTAACATTAAC 824

RESULT 7
LOCUS BE563259 846 bp mRNA linear EST 15-AUG-2000
DEFINITION 601335565F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3689369 5',
mRNA sequence.
ACCESSION BE563259
VERSION BE563259.1 GI:9807071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 846)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNCM384 Row: a Column: 18

High quality sequence stop: 709.
FEATURES
source
Location/Qualifiers
1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:3689369"
/clone_id="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOT7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."
BASE COUNT 172 a 252 c 266 g 156 t
ORIGIN
Query Match 25.8%; Score 698.6; DB 10; Length 846;
Best Local Similarity 96.9%; Pred. No. 1.3e-95;
Matches 776; Conservative 0; Mismatches 19; Indels 6; Gaps 6;
QY 474 CCTGCTGTAGTATTCTTACACGAGAGAGAGAGTGTAGCCAGATGAGAGCGTCTCT 533
Db 1 CCTGCTGTAGTATTCTTACACGAGAGAGAGAGTGTAGCCAGATGAGAGCGTCTCT 60
QY 534 GGGCGTGATTTCCCTGACAGCAGCTCCCGCAGCGACCCCTGAGGCGCTGCTCCCT 593
Db 61 GGGCGTGATTTCCCTGACAGCAGCTCCCGCAGCGACCCCTGAGGCGCTGCTCCCT 120
QY 594 CTGAGTGTACACCCAGGAGTGTACTTATGATGAGAGCGGGGCTTCAAGCTGAGCTGTG 653
Db 121 CTGAGTGTACACCCAGGAGTGTACTTATGATGAGAGCGGGGCTTCAAGCTGAGCTGTG 180
QY 654 GCAAGCCGGATCTTCAAGCCCATCTCATCCAGACATGTGTGGCCACACTCCAGTATT 713
Db 181 GCAAGCCGGATCTTCAAGCCCATCTCATCCAGACATGTGTGGCCACACTCCAGTATT 240
QY 714 GCACCAAGATGTGAGGAGCTCTAGGAGCGGCTTGTAGCCGGGTGGAGAGCCCTAC 773
Db 241 GCACCAAGATGTGAGGAGCTCTAGGAGCGGCTTGTAGCCGGGTGGAGAGCCCTAC 300
QY 774 CTGGGCCAGCCACTACCAAGAGAGACTGAACTCCAGACAGAGCTTCAATGATGTGAC 833
Db 301 CTGGGCCAGCCACTACCAAGAGAGACTGAACTCCAGACAGAGCTTCAATGATGTGAC 360
QY 834 GGTATGCGCCAGCTGAGTCTCTGCGGCTTCCAGCGCCGAGCTGCGGGTCTCTAGA 893
Db 361 GGTATGCGCCAGCTGAGTCTCTGCGGCTTCCAGCGCCGAGCTGCGGGTCTCTAGA 420
QY 894 ACAGAGCAGATGAGAGCGGCGATCGTGTGAGTGTGGAAGTGTGATGTGAGGA 953
Db 421 ACAGAGCAGATGAGAGCGGCGATCGTGTGAGTGTGGAAGTGTGATGTGAGGA 480
QY 954 CCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCC 1013
Db 481 CCTGAGAGTGTCACTTCCAAATAGATCCG-CAGGCTGTGAGAGTGGCGCTGGGGCTCC 539
QY 1014 CCTCAGCAGTATACGCTGACTTATGACACCAAGATGTGCTGCTGTGTGAGCAGCGGA 1073
Db 540 CCTCAGCAGTATACGCTGACTTATGACCA-CAGATGTGCTGTGTGAGCAGCGGA 598
QY 1074 CCGAGCTCCCGCATCTTCCCGCATCTTCCAGTGTGAGTGTGAGAGCGCAACACT 1133
Db 599 CCGAGCTCCCGCATCTTCCCGCATCTTCCAGTGTGAGTGTGAGAGCGCAACACT 657
QY 1134 GGAGAGCTGACAGAGAGAGAGGTCACCAATCTTTGAACAATGGCCCGGAGATTGACA 1193
Db 658 GGAGAGCTGACAGAGAGAGAGGTCACCAATCTTTGAACAATGG-CGCGGAGATTGA -AA 715
QY 1194 CTTTACCTTGAGCGCTTACCTACCAATGTGCGCTCTGGGATGAGAGTGGCGGCA 1253

Db 716 CTTTAACTGAGCGCTTCA-CTACCAATGTGGCTTCTGGATGAGAGTCCGCCA 774
OY 1254 GCTGCTGCCGACCTGAGAGA 1274
Db 775 CTGTGCGCTGTAGGACGCA 795

RESULT 8
BG340220
LOCUS BG340220 932 bp. mRNA linear EST 27-FEB-2001
DEFINITION 602438408F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556401 5',
mRNA sequence.
ACCESSION BG340220
VERSION BG340220.1 GI:13146647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
NIH-MGC http://mgc.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1258 row: P column: 02
High quality sequence stop: 630.
Location/Qualifiers
1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4556401"
/clone_lib="NIH-MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the Laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 160 a 284 c 326 g 162 t
ORIGIN

Query Match 25.7%; Score 695.6; DB 12; Length 932;
Best Local Similarity 93.0%; Pred. No. 3.5e-95;
Matches 794; Conservative 0; Mismatches 54; Indels 6; Gaps 6;

OY 52 GGGCGTGGCCGGTGCAGCCAGGTGTCGCGGCTGCTCCATGGCCCTGTCACAGTG 111
Db 1 GGGCGTGGCCGGTGCAGCCAGGTGTCGCGGCTGCTCCATGGCCCTGTCACAGTG 60
OY 112 AGCCCTTGGCCCGGGGAGGCGGCGCTCCAGCCCGTGGGGCCCTGGAGCAGCGGCTC 171
Db 61 AGCCCTTGGCCCGGGGAGGCGGCGCTCCAGCCCGTGGGGCCCTGGAGCAGCGGCTC 120
OY 172 CAGCAAGAGATGCACTCAGCAAGCAGAGACTTGGCGGTGCTCCGTGGGCTGCTCTG 231
Db 121 CAGCAAGAGATGCACTCAGCAAGCAGAGACTTGGCGGTGCTCCGTGGGCTGCTCTG 180
OY 232 GAGCTGCAAGATGAGAGGAGCAATGATGAGCAGAGAGCCAGTTCTGAGCCAACAGAG 291

Db 181 GAGCTGCAAGATGAGAGGAGCAATGATGATGAGCAGAGAGCCAGTTCTGAGCCAACAGAG 240
OY 292 AAGGCCCGGAGTGAAGAGAGCTCCAGCGGGGACAGACAGACTTGGGCAAGATCCAG 351
Db 241 AAGGCCCGGAGTGAAGAGAGCTCCAGCGGGGACAGACAGACTTGGGCAAGATCCAG 300
OY 352 AGTCCCAAGAAACAGAGAGAGCAGAGAGCAGCACTGCACTTCATGATACAGCTGCTGAG 411
Db 301 AGTCCCAAGAAACAGAGAGAGCAGAGAGCAGCACTTCATGATGATACAGCTGCTGAG 360
OY 412 CCGCAGATGATATCCGCTGCGACGCCAGCTGAGAGCAGCCGCGCTCCCGGCTCCG 471
Db 361 CCGCAGATGATATCCGCTGCGACGCCAGCTGAGAGCAGCCGCGCTCCCGGCTCCG 419
OY 472 TACCTGCTGATGTTTCTACACAGAGAGAGAGAGTGTGAGCAGAGATGAGAGCGCTCTC 531
Db 420 TACCTGCTGATGTTTCTACACAGAGAGAGAGAGTGTGAGCAGAGATGAGAGCGCTCTC 479
OY 532 CTGGGCTGATTTCCCTGACAGAGCTCCGACCTGACCCCTGAGGCTGCTTGGCC 591
Db 480 CTGGGCTGATTTCCCTGACAGAGCTCCGACCTGACCCCTGAGGCTGCTTGGCC 539
OY 592 CTCTGAGTGAACACCAAGGTACTTATGATGAGAGAGCGGGGCTTACAGCTGACGTGCT 651
Db 540 CTCTGAGTGAACACCAAGGTACTTATGATGAGAGAGCGGGGCTTACAGCTGACGTGCT 599
OY 652 GGGCAAGCGCGGATCTCAAGCCCATCTCCATCCAGACCATGAGGCGCACACTCCA-GGT 710
Db 600 GGGCAAGCGCGGATCTTCAAG-CCATCTCCATCAAGACCATGAGGCGCACACTCCA-GGT 658
OY 711 ATTGCACCAAGCATGTGAAGCAGCTCTAGG-CAGCGGCTTGTACCGGGTGGCAGTGCC 769
Db 659 ATTGCACCAAGCATGTGAAGCAGCTCTAGGCGCGGGGCTTGTTCGGTGGGAGTGCC 718
OY 770 TCACCTGGGCGACCACTACAGAGAGACTGACTCCAGACAGAGCTTCATGAT 829
Db 719 TCAGCTGGG-CAGCACTTCCCGGGGAGGCTGAGCTGCC-ACAGAGCTCTTCATGACT 776
OY 830 GAGCGGCTATGGCGCGACCTGGAGTCTGCGGGCTGCCAGCGCGGCTGGCGGCTCT 889
Db 777 GCGCGGCTATGGCGAGCCTGGAGGTTCTGCGGCTCCAGCCAGCCTGGGGCTCGACA 836
OY 890 CAGACAGCAGCAGCAG 903
Db 837 CCGGGCATGAGAGAG 850

RESULT 9
BE869222
LOCUS BE869222
DEFINITION 601445030F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849342 5',
mRNA sequence.
ACCESSION BE869222
VERSION BE869222.1 GI:10317998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 726)
NIH-MGC http://mgc.nci.nih.gov/
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM956 row: K column: 07

Db 419 GATGCTGCTGCTGTGGGACACAGCGGACCGAGCCTCCGCACTTCTCCCCACCTTACCT 478
QY 1107 GGGCTCAGAGTGGAGCGGCAACCTGTGGAGAGCTGCAGAGAACAGGGTCAACCCACT 1166
Db 479 GGGCTCAGAGTGGAGCGGCAACCTGTGGAGAGCTGCAGAGAACAGGGTCAACCCACT 538
QY 1167 CTTGAACATGCGCCGGAGATTGACAACTTCAACCTGAGCGCTTCACTTACACAAATGT 1226
Db 539 CTTGAACATGCGCCGGAGATTGACAACTTCAACCTGAGCGCTTCACTTACACAAATGT 598
QY 1227 GCGGCTGTGGAGTGAAGAGTCCGCCAGCTGCTGCCGACCTGGAAGAGACACCCGCTT 1286
Db 539 GCG-CTGTGGAGTGAAGAGTCCGG-CCAGCTGCTGC--GACTGGAAGAGACACCCGCTT 654
QY 1287 CATTGA-GGCTGCAAGAGACAGGAGACCCAGCTGCTGCTGCACATGCAAGATGGGCTCA 1345
Db 655 CATTGAAGGCTTCAAGAGACAGGAGACCCAGCTGCTGCTGCACATGCAAGATGGGCTCA 714
QY 1346 GCGGCTCAGCGGACAGTGTGCTGCTATGCTCATGAGCAGTACGATGACCTGAGAC 1405
Db 715 GCGGCTCAGCGG-CAAGAGGCTGCTGCTATGCTCATGAGCAGTACGATGCGGCTGAGAC 773
QY 1406 AGGCTCTGCGC 1416
Db 774 GCGCTGGGAC 784

RESULT 11

Bg753072

LOCUS 60273241P1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875562 5',
DEFINITION mRNA sequence.

ACCESSION

Bg753072

VERSION

Bg753072.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 663)

AUTHORS

NIH-MGC http://mgi.ncl.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1757 row: b column: 11
High quality sequence stop: 663.

FEATURES

source

1. 663

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4875562"

/clone_lib="NIH_MGC_43"

/tissue_type="normal pigmented retinal epithelium"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. 1"

BASE COUNT

124 a 205 c 231 g 103 t

Query Match 24.5%; Score 662; DB 12; Length 663;
Best Local Similarity 100.0%; Pred. No. 4, 66-90;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 AGGACTGTCGCGGGGTTGAGGAAAGGGCGTCCCGGTGCAGCCAGGTGCTCGCG 84
Db 2 AGGACTGTCGCGGGGTTGAGGAAAGGGCGTCCCGGTGCAGCCAGGTGCTCGCG 61
QY 85 CCTGGCTCCATGGCCCTGGTCAAGTGAAGCCCTTCCCGCCGGGCGAGCGCCCTCCAG 144
Db 62 CCTGGCTCCATGGCCCTGGTCAAGTGAAGCCCTTCCCGCCGGGCGAGCGCTCCAG 121
QY 145 CCCGTGGGGCCCTGGGAGCAGCGGCTCCAGCAAGAGTGCAGCTCCAGCAAGAGCAGAC 204
Db 122 CCCGTGGGGCCCTGGGAGCAGCGGCTCCAGCAAGAGTGCAGCTCCAGCAAGAGCAGAC 181
QY 205 TTTGCGGTCTCCGTGGGCTGCTGCTGGAATGCAAGATGAGGAGCAATGATGATGA 264
Db 182 TTTGCGGTCTCCGTGGGCTGCTGCTGGAATGCAAGATGAGGAGCAATGATGATGA 241
QY 265 GCGAGGCGCCAGTTCTGAGCAACAGAGAGAGGCGCCGAGTGAAGAGAGAGCTCCAG 301
Db 242 GCGAGGCGCCAGTTCTGAGCAACAGAGAGAGGCGCCGAGTGAAGAGAGAGCTCCAG 301
QY 325 CAGACAGACTTGGGCAAGATCCAGAGTCCCAAGACAGAGAGCAGAGCGAC 384
Db 302 CAGACAGACTTGGGCAAGATCCAGAGTCCCAAGACAGAGAGCAGAGCGAC 361
QY 385 CTGACACTTATGATGATGAGCTGCTGAGGCGCAGAGATGACATCCGCTGAGCGACGCTG 444
Db 362 CTGACACTTATGATGATGAGCTGCTGAGGCGCAGAGATGACATCCGCTGAGCGACGCTG 421
QY 445 GAGGAGCGCCGCGCTCCCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
Db 422 GAGGAGCGCCGCGCTCCCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 505 GGTCTGAGCAGAGATGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564
Db 482 GGTCTGAGCAGAGATGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
QY 565 AGCTGACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 542 AGCTGACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 625 GAGGCGGCTTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 684
Db 602 GAGGCGGCTTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
QY 685 CA 686
Db 662 CA 663

RESULT 12

BM910493

LOCUS

BM910493

DEFINITION

AGCNCOURT_6703004 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5453997

ACCESSION

BM910493

VERSION

BM910493.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1004)

AUTHORS

NIH-MGC http://mgi.ncl.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI948 row: g column: 22
High quality sequence start: 725.

BASE COUNT	211 a	299 c	324 g	167 t	3 others
ORIGIN					

Query Match	24.1%;	Score 652.4;	DB 14;	Length 1004;
Best Local Similarity	80.5%;	Pred. No. 1e-88;		

QY	165	GGCGGTCCAGCGAAGAGATGACTCCACGCGAAGGACAGAGTTTCGGTGTCCGTGGGGC	224
Db	1	GGCGGTCCAGCGAAGAGATGACTCCACGCGAAGGACAGAGTTTCGGTGTCCGTGGGGC	60
QY	225	TGTCCTGGGACTGCAGATGGAGGGGACAATGATGATGCAGCAGAGGCCAGTTCTGAGCC	284
Db	61	TGTCCTGGGACTGCAGATGGAGGGGACAATGATGATGCAGCAGAGGCCAGTTCTGAGCC	120
QY	285	AACGAGAAAGGCCCCGAGTGGAGGAGGCTCCACGGGGACCAGACACATTCGGGCAAG	344
Db	121	AACGAGAAAGGCCCCGAGTGGAGGAGGCTCCACGGGGACCAGACACATTCGGGCAAG	180
QY	345	ATCCAGAGTCCCGCAGAAAGCAGAGAGGAGCAGAGGCACACCTGCACCTCATGTAGACGT	404
Db	181	ATCCAGAGTCCCGCAGAAAGCAGAGAGGAGCAGAGGCACACCTGCACCTCATGTAGACGT	240
QY	405	GCTAGAGCCGAGGATGATCATCCGCCCTTGGCAGGCCACTGAGAGCACCCGGGCTCCCCG	464
Db	241	GCTAGAGCCGAGGATGATCATCCGCCCTTGGCAGGCCACTGAGAGCACCCGGGCTCCCCG	300
QY	465	GCCTCCGACTCGTGTGTAGTTTCTACACGAAAGGAGAGGTGTGAGCCAGATGAGAC	524
Db	301	GCCTCCGACTCGTGTGTGTAGTTTCTACACGAAAGGAGAGGTGTGAGCCAGATGAGAC	360
QY	525	GGTCTCTCTGGGCGTGAATTTCCCTGACAGCAGCTCCCCAGCTGCACCTGGGCGTGGT	584
Db	361	GGTCTCTCTGGGCGTGAATTTCCCTGACAGCAGCTCCCCAGCTGCACCTGGGCGTGGT	420
QY	585	CTTCCGCCCTCGGAGGTACACGCCGAGGTACTTAGATGGAGAGCGGGGCTTACGCGTAC	644
Db	421	CTTCCGCCCTCGGAGGTACACGCCGAGGTACTTAGATGGAGAGCGGGGCTTACGCGTAC	480
QY	645	GTCGTGTGGGCAAGACCGGATCTTCAAGACCATCTCCATCCAGACCTGTGTGGGCACT	704
Db	481	GTCGTGTGGGCAAGACCGGATCTTCAAGACCATCTCCATCCAGACCTGTGTGGGCACT	530
QY	705	CCAGGATTTGCACCAAGCATGTGAGGACGCTTAGGCAAGGCGCCTTTACCGGGTGGCAG	764
Db	531	CCAGGATTTGCACCAAGCATGTGAGGACGCTTAGGCAAGGCGCCTTTACCGGGTGGCAG	590
QY	765	TGCCCCTCACCTGGGGCCAGCCACTTACGAGGAGAGACTAAGTCCGAAACAGAGCTGCCAA	824
Db	531	TGCCCCTCACCTGGGGCCAGCCACTTACGAGGAGAGACTAAGTCCGAAACAGAGCTGCCAA	530

QY	825	TGAATGACGGCATATGCGCCGACCTGGAGTCTGTGGCGCTCCAGAGCGGACCTGGGGG	884
Db	531	-----G	531
QY	885	GTCCTCAGAACAGACGACGATGAGCAGGCGATCCGTGAGCTGTGGAAAGTTTGA	944
Db	532	GTCCTCAGAAACAGAGCAGATGAGAGCGAGCGATCCGTGAGCTGTGGAAAGTTTGA	591
QY	945	TGTTAGTAGCTTGAAGAGTTCACCTTCCAAAGAGATCCGCGAGGCTCTGGAAGCTCGCCT	1004
Db	592	TGTTAGTAGCTTGAAGAGTTCACCTTCCAAAGAGATCCGCGAGGCTCTGGAAGCTCGCCT	651
QY	1005	GGGGCTCCCCCTCCAGCAGTAGACCTGACTTCATCGAACCAAGATGCTCTGCTGTGC	1064
Db	652	GGGGCTCCCCCTCCAGCAGTAGACCTGACTTCATCGAACCAAGATGCTCTGCTGTGC	711
QY	1065	ACAGCGGGACCGAGCGCTCCGCCATCTTCCCCACCTCTACCTTGGGCTCAGAGTGAACGC	1124
Db	712	ACACCGGGACCGAGCGCTCCGCCATCTTCCCCACCTCTACCTTGGGCTCAGAGTGAACGC	771
QY	1125	AGCAACCTGGAGAGCTGTGAGAGAAACAGGGTACCCACATCTTGAACAT--GGCGCGGG	1183
Db	772	AGCAACCTGGAGAGCTGTGAGAGAAACAGGGTACCCACATCTTGAACATGGCGCGGGG	831
QY	1184	AGATTGCAACTTCTACTCTGAGCGGTTCACTACCAACCAATGTGGCGCTCTGGATGAGG	1243
Db	832	AGATTGCAACTTCTACTCTGAGCGGTTCACTACCAACCAAGTGGCGCTTCTGGGGATGAG	891
QY	1244	AGTGGCG 1250	
Db	892	AGGAGTC 898	

```

RESULT 13
B6679981 LOCUS
DEFINITION B6679981 653 bp mRNA linear EST 01-MAY-2001
602626694F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4751758 5',
mRNA sequence.
ACCESSION B6679981
VERSION B6679981
KEYWORDS B6679981.1 GI:13911378
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 653)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@email.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM10608 row: k column: 23
High quality sequence stop: 650.
Location/Qualifiers
1..653
FEATURES
source
```

```
BASE COUNT      143 a      223 c      173 g      114 t
ORIGIN
Query Match      23.7%; Score 642; DB 12; Length 653;
Best Local Similarity 99.8%; Pred. No. 4.6e-87;
Matches 653; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1694 GAGCTTCACATGAAGAGACCTCTGAGCCCTCCACAGCTTGCAGAGACCAAGGAGAGC 1753
      |||||||
DB 1 GAGCTTCACATGAAGAGACCTCTGAGCCCTCCACAGCTTGCAGAGACCAAGGAGAGC 60

OY 1754 CAGCAGGTGACAGAGGGGGCCTCAGCCCTGAGTCCCGCAGTCAGTGGTTACCCCTC 1813
      |||||||
DB 61 CAGCAGGTGACAGAGGGGGCCTCAGCCCTGAGTCCCGCAGTCAGTGGTTACCCCTC 120

OY 1814 CAGGCAATGTCCTGTGTGGCCCAACCGGACCCAGGCTTCCAGAGACAGAGAGAGGAGAG 1873
      |||||||
DB 121 CA-GGCAATGTCCTGTGTGGCCCAACCGGACCCAGGCTTCCAGAGAGAGAGAGGAGAG 179

OY 1874 GGGCAGGGGACAGGAGAGCCCTGCTATTCCTACAGCCAGGTTCCGAGAGGTGGTGAGA 1933
      |||||||
DB 180 GGGCAGGGGACAGGAGAGCCCTGCTATTCCTACAGCCAGGTTCCGAGAGGTGGTGAGA 239

OY 1934 CAGGCGACGCTGACATGACAGTGGAGAGAGGGCGAGGCTGAGCCCTCACAATGCCAC 1993
      |||||||
DB 240 CAGGCGACGCTGACATGACAGTGGAGAGAGGGCGAGGCTGAGCCCTCACAATGCCAC 299

OY 1994 GCTCCCTTCGACACTGAGAGAGATCCACAACCTCTTGGAGAAAACCCCTCAGCTGTTCG 2053
      |||||||
DB 300 GCTCCCTTCGACACTGAGAGAGATCCACAACCTCTTGGAGAAAACCCCTCAGCTGTTCG 359

OY 2054 CGCACACATTCCTCTCAGCTCCGGCCCATCCGCTACATACAGCTCAGTCCCAACCCCT 2113
      |||||||
DB 360 CGCACACATTCCTCTCAGCTCCGGCCCATCCGCTACATACAGCTCAGTCCCAACCCCT 419

OY 2114 GTCACTACGGGCTCAGCTCCAGCCCTGTCACTACAGCCCTCAGCTCAGCTTAACT 2173
      |||||||
DB 420 GTCACTACGGGCTCAGCTCCAGCCCTGTCACTACAGCCCTCAGCTCAGCTTAACT 479

OY 2174 CCCAGGCCATGTCTGCTGTCCGTCGAAGGCTCAAGACTTCTTAAGTGTGGATGAGG 2233
      |||||||
DB 480 CCCAGGCCATGTCTGCTGTCCGTCGAAGGCTCAAGACTTCTTAAGTGTGGATGAGG 539

OY 2234 GACGTAAGGTACCTTTGGGGGCAACAGCAGCCCTAGTTTCATTCCTCACTAGCCCTCA 2293
      |||||||
DB 540 GACGTAAGGTACCTTTGGGGGCAACAGCAGCCCTAGTTTCATTCCTCACTAGCCCTCA 599

OY 2294 CACTCACCCTGTGGCAGGGAATGAAAGAGAGCTTCCCGTGCAGAAAAGGTCACG 2347
      |||||||
DB 600 CACTCACCCTGTGGCAGGGAATGAAAGAGAGCTTCCCGTGCAGAAAAGGTCACG 653

RESULT 14
BE737087      919 bp      mRNA      linear      EST 15-SEP-2000
LOCUS      601304852F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3639229 5',
DEFINITION      mRNA sequence.
ACCESSION      BE737087
VERSION      BE737087.1 GI:10151079
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE      1 (bases 1 to 919)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-remail.nih.gov
      Tissue Procurement: ATCC
      cDNA Library Preparation: Ling Hong/Rubln Laboratory
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
```

```
FEATURES
      source
      1..919
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:3639229"
      /clone_id="NIH_MGC_39"
      /tissue_type="adenocarcinoma"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: pancreas; Vector: pOTB; Site:1: XhoI;
      Site:2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCAGAG(G). Library constructed
      by Ling Hong in the laboratory of Gerald M. Rubin
      (University of California, Berkeley) using ZAP-cDNA
      synthesis kit (Stratagene) and Superscript II RT (Life
      Technologies)."
```

```
BASE COUNT      170 a      269 c      316 g      164 t
ORIGIN
Query Match      23.7%; Score 641.6; DB 12; Length 919;
Best Local Similarity 93.5%; Pred. No. 4.5e-87;
Matches 746; Conservative 0; Mismatches 44; Indels 8; Gaps 7;

OY 615 CTTGATGAGAGCGGGGCTCAGCGTGACGTCGTGGGCAAGACCGGATCTTCAAGCC 674
      |||||||
DB 41 CTTATATGAGAGCGGGG--TTCAAGTACGCTGTGGGCAAGACCGGATCTTCAAGCC 98

OY 675 CATCTCATCAGACCATGTGGGCGACACTCCAGTATTGCAACAGCATGTAGGAGCG 734
      |||||||
DB 99 CATCTCATCAGACCATGTGGGCGACACTCCAGGTA--TGACACCAAGCATGTAGGAGCG 157

OY 735 TCTTAGGAGCGGCTTGTACCGGGTGAGTGCCCTCACTGGGGCAGCCACTACAGGA 794
      |||||||
DB 158 TCTTAGGAGCGGCTTGTACCGGGTGAGTGCCCTCACTGGGGCAGCCACTACAGGA 217

OY 795 GAGACTGAATCCGAACAGAGCTGCCTCAATGATGAGAGGCTATGGGCGACCTGGAATC 854
      |||||||
DB 218 GAGACTGAATCCGAACAGAGCTGCCTCAATGATGAGAGGCTATGGGCGACCTGGAATC 277

OY 855 TCTGGGCTCCAGCGCCGAGCCTGGCGGGTCTCAGAACAGAGAGAGATGAGAGCG 914
      |||||||
DB 278 TCTGGGCTCCAGCGCCGAGCCTGGCGGGTCTCAGAACAGAGAGAGATGAGAGCG 337

OY 915 GATCCGTGTGAGCTGTGTGAAAGTGTGGATGTCACTGAGAGATGTCACTTCCA 974
      |||||||
DB 338 GATCCGTGTGAGCTGTGTGAAAGTGTGGATGTCACTGAGAGATGTCACTTCCA 397

OY 975 AGAGATCCGCGAGGCTGTGGAGCTCGGCTGGGGGTCCCGCTCCAGAGTACCGGACTT 1034
      |||||||
DB 398 AGAGATCCGCGAGGCTGTGGAGCTCGGCTGGGGGTCCCGCTCCAGAGTACCGGACTT 457

OY 1035 CATGCACAACAGATGTGCTGTGTGGACAGGCGGAGACCGAGCCCTCCGATTTCC 1094
      |||||||
DB 458 CATGCACAACAGATGTGCTGTGTGGACAGGCGGAGACCGAGCCCTCCGATTTCC 517

OY 1095 CCACCTTACTGGGCTGAGAGTGAACGCAAACTGTGAGAGAGCTGCAAGAGAAAG 1154
      |||||||
DB 518 CCACCTTACTGGGCTGAGAGTGAACGCAAACTGTGAGAGAGCTGCAAGAGAAAG 577

OY 1155 GGTCAACCAATCTTG--AACATGGCCCGGAGATGACAACTTTACCCCTGAGCGCTTCA 1213
      |||||||
DB 578 GGTCAACCAATCTTGTAACATGCGCGGAGATTGACAA--TTTACCCCTGAGCGCTTCA 636

OY 1214 CTTACACACA-ATGTGCGGCTGTGGAGTGAAGTGGGCCAGGCTGCGCACATGTAAG 1272
      |||||||
DB 637 CTTACACACAGATGTGCGGCTGTGGAGTGAAGTGGGCCAGGCTGCGCACATGTAAG 695
```

```

QY 1273 GAGACGACCGCTTCATTGAGCTGCAGAGACACAGGCGACCCACGCTGCTGTCACCT-G 1331
      |||
DB 696 GAGACGACCGCTTCATTGAGCTGCAGAGACACAGGCGACCCACGCTGCTGCTGCTG 755
QY 1332 CAAGATGGGCGTCAGCGGCTCAGCGGCGACAGTGTGCTATGCTATGATGACAGTACGA 1391
      |||
DB 756 CAGATGGGCGGCTCAGCGGCGGCGGCGGCGGCTGCTATGCTGCTGATGAGCAGTACGA 815
QY 1392 ATGCACCTGTGAGCAGCGC 1409
      |||
DB 816 ATGCCCTGTGTGCGCGCC 833

RESULT 15
BE907514 716 bp mRNA linear EST 20-OCT-2000
LOCUS BE907514
DEFINITION 601497473fl NIH_MGC_70 Homo sapiens CDNA clone IMAGE:3899418 5',
      mRNA sequence.
ACCESSION BE907514 GI:10401153
VERSION BE907514
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 716)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Life Technologies, Inc.
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/ULNL at:
      http://image.llnl.gov
      Plate: LRAM9697 row: a column: 19
      High quality sequence stop: 683.
      Location/Qualifiers
        1. 716
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="IMAGE:3899418"
          /clone_lib="NIH_MGC_70"
          /tissue_type="epithelioid carcinoma"
          /lab_host="DH10B (phage-resistant)"
          /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
          Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
          Average insert size 1.1 kb. library constructed by Life
          Technologies."
BASE COUNT 138 a 213 c 251 g 114 t
ORIGIN

Query Match 23.5%; Score 635.4; DB 12; Length 716;
Best Local Similarity 98.3%; Pred. No. 4.3e-86;
Matches 705; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

```

```

DB 180 GGCAGAGCTTTTCGCGTCCCTGGGCGTCTGTGGATGAGATGAGGGGCAATG 239
      |||
QY 257 ATGATCCACACAGGCGCACTTCAGGCGCAACAGAGAGGCGCCGAGTGGAGAGCTCC 316
      |||
DB 240 ATGATGACAGAGGCGCACTTCAGGCGCAACAGAGAGGCGCCGAGTGGAGAGCTCC 299
      |||
QY 317 ACGGGACACAGAGACTTCGAGGCAAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAG 376
      |||
DB 300 ACGGGACACAGAGACTTCGAGGCAAGATCCAGAGTCCCGAGAGAGAGAGAGAGAG 359
      |||
QY 377 GGCAGACCTGACCTCATGTGACAGCTGCTAGGCGCGAGATGACATCCGCTGGCAG 436
      |||
DB 360 GGCAGACCTGACCTCATGTGACAGCTGCTAGGCGCGAGATGACATCCGCTGGCAG 419
      |||
QY 437 CCAGCTGAGGACACCGCGGCGCTCCCGGCTCCGCTACCTGCTGGAGTTCATACAGAG 496
      |||
DB 420 CCAGCTGAGGACACCGCGGCGCTCCCGGCTCCGCTACCTGCTGGAGTTCATACAGAG 479
      |||
QY 497 AAGGAGAGGCTGAGCGCAGATGAGACGCGCTCCGCGCGTGAATTTCCCTGACAGCA 556
      |||
DB 480 AAGGAGAGGCTGAGCGCAGATGAGACGCGCTCCGCGCGTGAATTTCCCTGACAGCA 539
      |||
QY 557 GCTCCCGCAGCTGACCGCTGGGCGCTGCTTCCCTCTGGAGTGACACCGCAGGTGACT 616
      |||
DB 540 GCTCCCGCAGCTGACCGCTGGGCGCTGCTTCCCTCTGGAGTGACACCGCAGGTGACT 599
      |||
QY 617 TAGATGAGACGCGGCGCTGACGCTGAGTGGGCAAGCGGATCTTCAAGCGCA 676
      |||
DB 600 TAGATGAGACGCGGCGCTGACGCTGAGTGGGCAAGCGGATCTTCAAGCGCA 657
      |||
QY 677 TCTCATCCAGACCATGTGGGCGCAGACTCCAGATATTGACCAAGCATGTGAGGAG 733
      |||
DB 658 TCTCATCCAGACCATGTGGGCGCAGACTCCAGATATTGACCAAGCATGTGAGGAG 712
      |||

```

Search completed: April 11, 2003, 23:48:52
 Job time : 7374.79 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2003, 00:00:35 ; Search time 156.754 Seconds
(without alignments)
5290.174 Million cell updates/sec

Title: US-09-761-640-1

Perfect score: 2704

Sequence: 1 cgtctctctctgtctctgcyg.....aaaaaaaaaaaaaaaaaaaaa 2704

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104.4	3.9	1390	US-09-704-139-1	Sequence 1, Appli
2	103.4	3.8	531	US-09-704-139-3	Sequence 3, Appli
3	88	3.3	1238	US-08-530-290-11	Sequence 11, Appli
4	78.4	2.9	1987	US-08-990-379-1	Sequence 1, Appli
5	78.4	2.9	1993	US-08-990-379-2	Sequence 2, Appli
6	61.6	2.3	1729	US-09-045-973-6	Sequence 6, Appli
7	60	2.2	7218	US-08-232-463-14	Sequence 14, Appli
8	56.2	2.1	594	US-09-163-833-3	Sequence 3, Appli
9	56.2	2.1	912	US-09-163-833-1	Sequence 1, Appli
10	55.4	2.0	1691	US-09-013-881-12	Sequence 12, Appli
11	52.6	1.9	216	US-08-530-290-9	Sequence 9, Appli
12	50.2	1.9	861	US-07-988-273-1	Sequence 9, Appli
13	50.2	1.9	861	PCR-US93-12019-1	Sequence 1, Appli
14	48.6	1.8	477	US-09-135-994-1	Sequence 1, Appli
15	48.2	1.8	237	US-08-530-290-8	Sequence 8, Appli
16	46.8	1.7	1926	US-09-249-585A-2	Sequence 2, Appli
17	46.8	1.7	2580	US-09-050-863-2	Sequence 2, Appli
18	46.8	1.7	2580	US-09-359-081-2	Sequence 2, Appli
19	46.8	1.7	5452	US-09-130-114-1	Sequence 2, Appli
20	46.8	1.7	9600	US-08-910-647-1	Sequence 1, Appli
21	46.8	1.7	9600	US-09-620-925-1	Sequence 1, Appli
22	46.8	1.7	10596	US-07-884-811-15	Sequence 15, Appli
23	46.8	1.7	10596	US-07-885-971-15	Sequence 15, Appli
24	46.8	1.7	10596	US-08-087-783A-15	Sequence 15, Appli
25	46.8	1.7	10596	US-08-194-088B-15	Sequence 15, Appli
26	46.8	1.7	10596	US-08-194-087-15	Sequence 15, Appli
27	46.8	1.7	10596	PCR-US93-04648-15	Sequence 15, Appli

C 28	46.2	1.7	1926	US-09-249-585A-2	Sequence 2, Appli
C 29	46.2	1.7	2580	US-09-050-863-2	Sequence 2, Appli
C 30	46.2	1.7	2580	US-09-359-081-2	Sequence 2, Appli
C 31	46.2	1.7	5452	US-09-130-114-1	Sequence 1, Appli
C 32	46.2	1.7	9600	US-08-910-647-1	Sequence 1, Appli
C 33	46.2	1.7	9600	US-09-620-925-1	Sequence 1, Appli
C 34	46.2	1.7	10596	US-07-884-811-15	Sequence 15, Appli
C 35	46.2	1.7	10596	US-07-885-971-15	Sequence 15, Appli
C 36	46.2	1.7	10596	US-08-087-783A-15	Sequence 15, Appli
C 37	46.2	1.7	10596	US-08-194-088B-15	Sequence 15, Appli
C 38	46.2	1.7	10596	US-08-194-087-15	Sequence 15, Appli
C 39	46.2	1.7	10596	PCR-US93-04648-15	Sequence 15, Appli
C 40	45.6	1.7	2214	US-08-864-038A-1	Sequence 1, Appli
C 41	45.6	1.7	3331	US-08-864-038A-2	Sequence 2, Appli
C 42	45.6	1.7	3331	US-08-864-038A-4	Sequence 4, Appli
C 43	45.4	1.7	1358	US-09-013-881-11	Sequence 11, Appli
C 44	44	1.6	7218	US-08-232-463-14	Sequence 14, Appli
C 45	43.8	1.6	35828	US-09-449-218D-17	Sequence 17, Appli

ALIGNMENTS

```

RESULT 1
US-09-704-139-1
; Sequence 1, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREOF
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (329)..(859)
; NAME/KEY: misc.feature
; LOCATION: (1)..(1390)
; OTHER INFORMATION: n = a, t, g, or c
US-09-704-139-1

```

Query Match 3.9%: Score 104.4; DB 4; Length 1390;
Best Local Similarity 51.3%: Pred No. 7.2e-15;
Matches 243; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

```

QY 1084 CGCATCTTCCCGCCACCTTACCTGAGCTGAGTGAAGCAAGCAAACTGGAGAGCTG 1143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 CGCGTGCAGCCCTCACTCTTCTTCGAGAGCGCGGAGCGCGGAGGAGGAGGAGCTG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1144 CAGAGGAGAGAGGATGACCCATCTTGAACATGAGCCCGGAGATTGACAACTTCACTT 1203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 GCGCGGCGGAGGATGACGCTGTGCGTCAACGTCCTCCGACAGCCCGCGCGCGG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1204 GAGCGCTTACCTACACAAATGCGCGCTCTGGGATGAGAGAGCGCGCCAGCTGCGCG 1263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CCGCGCGTGCAGAGCTGCGCGCTGCCCGTTCGACAGACCGCGCTGAGACCTGCTGCGG 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1264 CACTGGAAGAGAGAGCGACCGCTTCACTTGAAGCTGCAAGAGCACAGGCGACCCAGCTGCTG 1323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 CACCTGAGACCCACGAGCGCGCCGATGAGAGCGCGCGGCTGCGCGCGCGCGCTGCTGCTG 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1324 GTCCACTGCAAGATGAGCGCTGAGCGCGCTGAGCGCGCACAGTCTGCTATGCTATGAAG 1383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 GTCTACTGCAAGAGAGCGCGCGCGCTGCGCGCTGCGCGCTGCTGACGCGCTACTCATCTGCG 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1384 CAGTAGAATGACGCTGAGCAGGCGCCGACGTCACGAGCTCCGCCATCCG 1443
    || || || || || || || || || || || || || || || || || || ||
Db 689 CACCGCGGCTCAGCGCTGAGCGCTTCCAGATGTGAAGAGCGCTCCCGGAGCA 748
QY 1444 CGCCCAACCGCTGCTCTGCGGACGTCGAGATCTACGAGGCGATCTGAGCGGCA 1503
    || || || || || || || || || || || || || || || || || || ||
Db 749 GAACGGAACCGCGCTTCTGCTCAGCTCCAGAGTATGAGAGAGGCGCTCCAGGCGCCAG 808
QY 1504 ACCTAGAGGTGTGGGAGAGAGAGTGTAGCATGGAAGAGAGCGCAGC 1557
    || || || || || || || || || || || || || || || || || || ||
Db 809 TCCTGCTCAGAGGAGAGCGCCCGACGCTTAGGTTGGCGCTGAGGCTGAAGC 862

```

```

RESULT 2
US-09-704-139-3
; Sequence 3, Application US/09704139
; Patent No. 6420153
;
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; PRIORITY FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIORITY FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 3
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-704-139-3

```

```

Query Match          3.8%; Score 103.4; DB 4; Length 531;
Best Local Similarity 52.7%; Pred. No. 8,1e-15;
Matches 224; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

```

QY 1084 CGCATCTTCCCGACCTTACCTGAGCTGAGTGAAGCAGCAAACTGAGAGACTG 1143
    || || || || || || || || || || || || || || || || || || ||
Db 61 CGGCTGCGCGCTCAGCTTCTCTCGGAGCGCGCAGAGCGCGGCGCGAGAGAGCTG 120
QY 1144 CAGAGAACAGGTCACCCACATCTTGAACATGGCCCGGAGATGACAACTTACCT 1203
    || || || || || || || || || || || || || || || || || || ||
Db 121 GCGCGCGCGGAGTACAGCTGTGCTAACGTCTCCCGCAGCAGCCCGCGCGCGG 180
QY 1204 GAGCGCTTACCTACCAATGTGCGCTCTGGGATGAGAGATGCGCCAGCTGCGG 1263
    || || || || || || || || || || || || || || || || || || ||
Db 181 CCGCGCTGCGAGAGCTGCGCGCTGCTTTCGACGAGACCCGCGCTGAGACCTGCGG 240
QY 1264 CACTGGAAGAGAGCAGCCGCTTATGAGGCTGCAAGAGACAGAGGACCCACGTGCTG 1323
    || || || || || || || || || || || || || || || || || || ||
Db 241 CACCTGAGCGCCACGTGCGCGCGCATGAGAGCGCGGCTGCGCGCGCGCTGCTA 300
QY 1324 GTTCATGCAAGATGGGCGTCACCGCTCAGCGGCGCACATGCTGTGCTATGCAATGAAG 1383
    || || || || || || || || || || || || || || || || || || ||
Db 301 GTTCTATGCAAGAGCGCGCGCGCTGCGCGCTGCTGCAACCGGTACTTCTGCGG 360
QY 1384 CATTAGCAATGCAAGCTGAGAGCGCCCTGCGCACGTCGAGAGAGCTCCGCGCATGCGC 1443
    || || || || || || || || || || || || || || || || || || ||
Db 361 CACCGGCGCTCAGCTGCGGAGAGCGCTTCCAGATGTTGAAGAGCGCTCCCGGTAGCA 420
QY 1444 CGCCCAACCGCTGCTCTGCGGACGTCGAGATCTACGAGGCGATCTGAGCGGCGAGA 1503
    || || || || || || || || || || || || || || || || || || ||
Db 421 GAACGGAACCGCGCTTCTGCTCAGCTCCAGAGTATGAGAGAGGCGCTCCAGGCGCCAG 480
QY 1504 ACCTG 1508
    || || || || || || || || || || || || || || || || || || ||
Db 481 TCCTG 485

```

```

RESULT 3
US-08-530-290-11
; Sequence 11, Application US/08530290

```

```

; Patent No. 5958721
;
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-000000US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-530-290-11

```

```

Query Match          3.3%; Score 88; DB 2; Length 1238;
Best Local Similarity 50.5%; Pred. No. 4,1e-11;
Matches 214; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

```

```

QY 1087 ATCTTCCCGACCTTACCTGAGGCTCAGAGTGAAGCAGCAAACTGAGAGCTGAC 1146
    || || || || || || || || || || || || || || || || || || ||
Db 648 ATCTTCCCTCTCTTACTGCTGCGAGTGTCTACCATGCTGCCGAGAGACATGCTGGAG 707
QY 1147 AGAAGAGGCTACCCACATCTTGAACATGCGCGGAGATTGACAACTTACCTTGAG 1206
    || || || || || || || || || || || || || || || || || || ||
Db 708 GCCCTGGCATCACGGGCTGTGTAATGTCTCTCGAGACTGCCCAACCACTTTGAAGA 767
QY 1207 CGCTTACCTACCAATGTGCGCTCTGGGATGAGAGAGTGGCCCGACAGCTGCGCGGAC 1266
    || || || || || || || || || || || || || || || || || || ||
Db 768 CACTATCATGACAAGTGCATCCAGTGAAGATACCAACAGGCGACATCAGCTCTG 827
QY 1267 TGAAGAGAGAGCAGCGCTTATGAGGCTGCAAGAGACACAGGCAACCCAGTGTGTC 1326
    || || || || || || || || || || || || || || || || || || ||
Db 828 TTCAATGGAAGCATAGAGTATCATGATCCCTGAAGAGACTGCGCTGCGCGCTGCTG 887
QY 1327 CACTGCAAGATGGCGCTCAGCGCTCAGCGGCGACAGAGTGTGCTTATGCAATGAAGCAG 1386
    || || || || || || || || || || || || || || || || || || ||
Db 888 CACTGCGAGGCGGACATCTGCGGTGCGGCACCATCTGCTGCTTACTGATGATGAAG 947

```

OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGATCCGGCCATCGCCCGC 1446
 DB 948 AACGGGTGAGGCTGAGAGAGGCCCTTCGATGCTTAAGACAGCGCCGACGATCATCTCG 1007
 OY 1447 CCCAACCTGCTTCTCTCGCCAGCTGCAGATCTACACGGGCATCTGACGGCCAGAAC 1506
 DB 1008 CCCAATTGAGCTTCATGCGGAGCTGCTGAGTTCAGTCCAGAGTCTGCGCCAGCTCC 1067
 OY 1507 TGAG 1510
 DB 1068 TGTG 1071

RESULT 4

US-08-990-379-1
 : Sequence 1, Application US/08990379
 : Patent No. 5998188
 : GENERAL INFORMATION:
 : APPLICANT: Stork, Philip J
 : APPLICANT: Misra-Press, Anita
 : TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
 : TITLE OF INVENTION: Their Biologically Active Expression Products
 : FILE REFERENCE: 4104-000322USA
 : CURRENT APPLICATION NUMBER: US/08/990, 379
 : CURRENT FILING DATE: 1997-12-15
 : EARLIER APPLICATION NUMBER: PCT/US96/10402
 : EARLIER FILING DATE: 1996-06-14
 : EARLIER APPLICATION NUMBER: 60/000, 263
 : EARLIER FILING DATE: 1995-06-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 1
 : LENGTH: 1987
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-08-990-379-1

Query Match 2.9%; Score 78.4; DB 2; Length 1987;
 Best Local Similarity 49.1%; Pred. No. 8e-09;
 Matches 208; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

OY 1087 ATCTTCCCCACCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAG 1146
 DB 969 ATCTTCTTCTCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAG 1028
 OY 1147 AGAAGACGGGTACCCACATCTTGAACATGGCCCGGAGATTGAACCTTCAACCTGAG 1206
 DB 1029 GCCCTGGGATACGGCTCTACTGATGTCTCTCAGACTGCCCAATACCTTTGAGGGA 1088
 OY 1207 CGCTTACCTACACATGATGGGCTCTGAGTGAAGAGTGGCCAGCTGCTGCGCAC 1266
 DB 1089 CATTCACGATACAGTGCATCCCGGTAGAGATTAACCAAGGCTGACATGCTCTGCG 1148
 OY 1267 TGAAGAGAGACGACCGCTTCACTTGAAGCTGACAGACACAGGAGCCAGCTGCTGTC 1326
 DB 1149 TTCATGAGAGCCATGATATAGACGATGAGGAGCTGCGGAGGCGAGTGTGCTT 1208
 OY 1327 CACTGACGATGGGCTGAGCGGCTCAGGCGCACAGTGTGCTATGCTATGAGAGG 1386
 DB 1209 CACTGCGAGCGCGGATCTTAGATACAGCACCATCTGCTGCTACCTGATGAGAG 1268
 OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGCTTCGGCCATCGCCGC 1446
 DB 1269 AAACGGGTGAGGCTGAGAGAGGCTTTCGAGTTCGATCAAGACAGCGCGTAGACATCATCTCG 1328
 OY 1447 CCCAACCTGCTTCTCTCGCCAGCTGCAGATCTACACGGGCATCTGACGGCCAGAAC 1506
 DB 1329 CCCAATTGAGCTTCATGCGGAGCTGCTGAGTTCAGTCTCAGGTGCTCACACAGCTCC 1388
 OY 1507 TGAG 1510
 DB 1389 TGTG 1392

RESULT 5
 US-08-990-379-2
 : Sequence 2, Application US/08990379
 : Patent No. 5998188
 : GENERAL INFORMATION:
 : APPLICANT: Stork, Philip J
 : APPLICANT: Misra-Press, Anita
 : TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
 : TITLE OF INVENTION: Their Biologically Active Expression Products
 : FILE REFERENCE: 4104-000322USA
 : CURRENT APPLICATION NUMBER: US/08/990, 379
 : CURRENT FILING DATE: 1997-12-15
 : EARLIER APPLICATION NUMBER: PCT/US96/10402
 : EARLIER FILING DATE: 1996-06-14
 : EARLIER APPLICATION NUMBER: 60/000, 263
 : EARLIER FILING DATE: 1995-06-16
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 1993
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-08-990-379-2

Query Match 2.9%; Score 78.4; DB 2; Length 1993;
 Best Local Similarity 49.1%; Pred. No. 8.1e-09;
 Matches 208; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

OY 1087 ATCTTCCCCACCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAG 1146
 DB 975 ATCTTCTTCTCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAG 1034
 OY 1147 AGAAGACGGGTACCCACATCTTGAACATGGCCCGGAGATTGAACCTTCAACCTGAG 1206
 DB 1035 GCCCTGGGATACGGCTCTACTGATGTCTCTCAGACTGCCCAATACCTTTGAGGGA 1094
 OY 1207 CGCTTACCTACACATGATGGGCTCTGAGTGAAGAGTGGCCAGCTGCTGCGCAC 1266
 DB 1095 CATTCACGATACAGTGCATCCCGGTAGAGATTAACCAAGGCTGACATGCTCTGCG 1154
 OY 1267 TGAAGAGAGACGACCGCTTCACTTGAAGCTGACAGACACAGGAGCCAGCTGTGCTC 1326
 DB 1155 TTCATGAGAGCCATGATATAGACGATGAGGAGTGTGCTGAGTGTGCTT 1214
 OY 1327 CACTGACGATGGGCTGAGCGGCTCAGAGTGAAGAGTGGCCAGCTGCTATGAGAGG 1386
 DB 1215 CACTGCGAGCGCGGATCTTAGATACGACCATCTGCTGCTTACCTGATGAGAG 1274
 OY 1387 TAGCAATGACGCTTGAGACAGGCCCTGCGCCACGTGACAGAGCTTCGGCCATCGCCGC 1446
 DB 1275 AAACGGGTGAGGCTGAGAGAGGCTTTCGAGTTCGATCAAGACAGCGCGTAGACATCATCTCG 1334
 OY 1447 CCCAACCTGCTTCTCTCGCCAGCTGCAGATCTACACGGGCATCTGACGGCCAGAAC 1506
 DB 1335 CCCAATTGAGCTTCATGCGGAGCTGCTGAGTTCGAGTCTCAGGTGCTCACACAGCTCC 1394
 OY 1507 TGAG 1510
 DB 1395 TGTG 1398

RESULT 6

US-09-045-973-6
 : Sequence 6, Application US/09045973
 : Patent No. 6165767
 : GENERAL INFORMATION:
 : APPLICANT: Lal, Preeti
 : APPLICANT: Yue, Henry
 : APPLICANT: Corley, Neil C.
 : APPLICANT: Guejter, Karl J.
 : APPLICANT: Baughn, Mariah
 : TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES

```

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTW0116
CLONE: 3041794

```

Query Match	Similarity	2.3%	Score	61.6	DB	Length	1729
Best Local	Similarity	46.8%	Pred	No.5	6e-05		
Matches	193	Conservative	0	Mismatches	219	Indels	0
							Gaps
							0
QY	1082	CCCCGACTCTTCCGCCACCTCTACCTGGGCTCGAGGTGAAGCAACGACCAACCTGGAGAGAC	1141				
Db	574	CTCAATATCACTCTCTCTATTCCTGGGAGAGGAGTGCTGGCGCTTCATTCGACACTCC	633				
QY	1142	TGCAGAGGAAACGGGTGCACCCCACTCTTGAAATAGCCCGGGAGATTGCACACTTTACC	1201				
Db	634	TCCAGGCTCGTGGATCACCTGCATTTGTTAAATGCTACCAATGTGATCCCTAATTTTAACT	693				
QY	1202	CTGAGCGGCTTCACTACCACTATGCGCGCTTGGGATGAGGAGTGGCGGACCTGCTGC	1261				
Db	694	GGCCCCAATTTTAGTATGTGTTAAAGTBCCTCTGTGGCTACATGCCGCAATGCCCATTTGGAC	753				
QY	1262	CGCACTGGAAGAGAGACGCCCTTCATTATGAGGCTGCAAGAGACAGGGGACCCACCTGCG	1321				
Db	754	TCTACTTTTGACACCGTGCGCTGCAACAGATCCACAGTGTGAGACAGAGCAACGGGGCCACTT	813				
QY	1322	TGTCCTACTGCAAGATGGCGCTCAGCCGCTCAGCGGCCACAGTGTCTGGCTATGCCATGA	1381				
Db	814	TGGTCACACTGCTCTGCAGGGGTGAGCGGCTCACGCCAGCGCTGTGTATCGGTTACTTGTA	873				
QY	1382	AGCAGTAGCAATGACAGCCCTGGAGAGCGCCCTGGCCACGTCGACGAGAGCTCCGGCCCATC	1441				
Db	874	AATTTCCACACAGTGTBCTCTGCTGGAGGCTTACACACTGGGTGAAAGGCCCGGCGACTGTCA	933				
QY	1442	CCCGCCCAACCTTGCTTCTCGCGCAGCTGCAGATCTACCAAGGGCACTCT	1493				
Db	934	TCAGGCCCAACGTAAGGCTTCTTGAGGCAACTGATAGACTACGAGGCCCACT	985				

```

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-Fls
US-08-232-463-14

```

[illegible]

[illegible]

```

US-08-530-290-9
; Sequence 9, Application US/08530290
; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-00000005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEO ID NO.: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-530-290-9

Query Match 1.9%; Score 52.6; DB 2; Length 216;
Best Local Similarity 53.7%; Pred. No. 0.0028;
Matches 109; Conservative 0; Mismatches 94; Indels 0; Gaps

QY 1262 CGCAGTGGAGAGAGCGACCGCTTCATTGAGCGTCGACAGACACAGGCGACCACTGTC 13221
14 CCCACTTTCACAAACCAATTGATTATTTATGTACTGCGTCAGGGAAGAGAGGCAAGGTC 73

```

OY	1332	TGCTCCACTGCAGATGGCGGTACGCCGCTCAAGGGSCACAAGTGGCGGCTTAGTGCCATGA	1381
Dd	74	TAGTCACCTGTGAAGGCTGGGGTTCTCAGAGGTACACCACCATCTTGCAATGGGGTACTCTCATAGA	133
OY	1382	AGCAGTTCAGATGAGCGCTGGACAGACGCCCTGGCGCACGTGCAGAGAGCTCCGGGCCATCG	1441
Dd	134	AGACCAAGCACTTCCGCTTAAGSAGGCCCTTGACATCGTCAAGCAGAGSAGSAGAACGTGA	193
OY	1442	CCGCCCCCAACCCCTGGCTTCCTG	1464
Dd	194	TCCTCCCAACTTGGCTTATG	216

QY	1434	GC	CC	AT	GC	CC	GC	CC	CA	AC	CT	GT	GC	CC	CA	GC	CT	1472
QY	1374	TG	CA	TG	CA	TG	CA	TG	CA	TG	CA	TG	CA	TG	CA	TG	CA	1433
Db	382	CG	CG	GT	CG	GT	CG	GT	CG	GT	CG	GT	CG	GT	CG	GT	CG	441
QY	442	CG	CG	AT	GC	CC	GC	CC	CA	AC	CT	GT	GC	CC	CA	GC	CT	501
QY	1434	GC	CC	AT	GC	CC	GC	CC	CA	AC	CT	GT	GC	CC	CA	GC	CT	1472

Db 502 TGAGATCG---GCCCAACGATGCTTCTGCGCCAGCT 537

RESULT 13

PCT-US93-12019-1

Sequence 1, Application PC/TUS9312019
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12019
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,273
FILING DATE: 14-DEC-1992
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 29..586
PCT-US93-12019-1

Query Match 1.9%; Score 50.2; DB 5; Length 861;
Best Local Similarity 61.6%; Pred. No. 0.018;
Matches 98; Conservative 0; Mismatches 58; Indels 3; Gaps 1;

Qy 1314 CCACGTGCTGCTCACTGATGAGCGGCTGAGCGGCGCACAGTGTGCGCTA 1373
Db 382 CGGGTGTCTGCTCACTGCGGGAAGTTATAGCCGCTCCCAAGCGTATGCTGCTA 441
Qy 1374 TGCATGAAGAGTACGATGAGCGGCTGAGCGGCGCACAGTGTGCGCTGCG 1433
Db 442 CCTCATGATGCGGAGAGATGAGTCAATGCTGCGCTGAGCATGCTGAGCGAGAACG 501
Qy 1434 GCCATCGCGCGCCCAACCCCTGCTTCTGCGCCAGCT 1472
Db 502 TGAGATCG---GCCCAACGATGCTTCTGCGCCAGCT 537

RESULT 14

US-09-135-994-1

Sequence 1, Application US/09135994A
Patent No. 6280938
GENERAL INFORMATION:
APPLICANT: Ranum et al.
TITLE OF INVENTION: SCAL GENE AND METHODS OF USE
FILE REFERENCE: University of Minnesota
CURRENT APPLICATION NUMBER: US/09/135,994A
EARLIER FILING DATE: 1998-08-18
EARLIER APPLICATION NUMBER: 60/056,170
EARLIER FILING DATE: 1997-08-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 477
TYPE: DNA
ORGANISM: Homo sapiens
US-09-135-994-1

Query Match 1.8%; Score 48.6; DB 4; Length 477;
Best Local Similarity 51.1%; Pred. No. 0.032;

Matches 114; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 258 TGATGACAGAGAGGCGGCTTGTAGCCACAGAGAGCCCGGATGAGAGAGCTCCA 317
Db 30 TGTAGAGCGGGAAGATATTCGAGCGGCGCCGGATGACATGAGGGGAGAGCGCGCG 89
Qy 318 CGGGACAGACAGACTTGGGCAAGATCCAGACTCCCGAAGACAGAGAGAGCAGAG 377
Db 90 CCGGCGGCGGCGGCGGAGAGAGCGGCGCGCGCGGAGAGAGAGAGAGAGAGAGCA 149
Qy 378 GCAGACCTGACACTTATGTTACAGTGTGAGCGCGGAGATGACATCCGCTGCGACG 437
Db 150 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 209
Qy 438 CCAGCTGAGAGCAGCCCGGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 210 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 252

RESULT 15

US-08-530-290-8

Sequence 8, Application US/08530290
Patent No. 5958721
GENERAL INFORMATION:
APPLICANT: Marshall, Christopher John
APPLICANT: Ashworth, Alan
APPLICANT: Hughes, David Anthony
TITLE OF INVENTION: Methods for Screening of Substances for
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,290
FILING DATE: 14-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB94/00694
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9402573.1
FILING DATE: 10-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9307250.2
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084611-00000005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-530-290-8

Query Match 1.8%; Score 48.2; DB 2; Length 237;
Best Local Similarity 50.2%; Pred. No. 0.03;

Matches 119; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```

Oy 1087 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACGCAACCTGGAGAGTGCAG 1146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1  ATCTTGGCCCTTACCTGTTCCTGGGCTGCACTCACTGCTGAGACCTGGAGGGCTGCAG 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1147 AGGACAGGGGTCCACCATCTTGAACATGGCCCGGAGATTGACACTTCTACCCCTGAG 1206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   61 GCGTGTGGGATACAGCCGCTCTCAACGCTGCCGCCAGCTGCCCAACCACTTTGAGGGC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1207 CGCTTCACTTACCAATATGTGGGCTCTGGGATGAGAGTGGGCCAGCTGCTGCCGCAC 1266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   121 CTTTCCGCTACAGAGATATCCCTGTGGAGGACCAACAGATGTGGAGATCAGTGCCTGG 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 1267 TGGAGGAGAGCAGCCGCTTCATTTAGGCTGCCAGAGCACAGGCAACCCAGCTGCTG 1323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   181 TTCCAGGAGGCCATAGGCTTCACTGTGGTGAAGAACAGCGGAGCCGGGTGCTG 237
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: April 11, 2003, 23:54:13
 Job time : 188.754 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:54:09 ; Search time 12816.2 Seconds

(without alignments)
5767.767 Million cell updates/sec

Title: US-09-761-640-3

Perfect score: 2540

Sequence: 1 ccgcgtccctcgcggtccagcag.....aaaaaaaaaaaaaaaaaaaaa 2540

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pl: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_om: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_man: *
37: em_htg_vtl: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	84.7	2808	9 AK074432	AK074432 Homo sapi
2	2115	83.3	2781	6 AK099939	AK099939 Sequence
3	1950.4	76.0	2905	9 AK094226	AK094226 Homo sapi
4	1641	64.6	1905	9 BC004176	BC004176 Homo sapi
5	1641	64.6	1905	9 BC004210	BC004210 Homo sapi
6	1345.6	53.0	1980	6 AK451343	AK451343 Sequence
7	1303.6	51.3	1755	6 AK398768	AK398768 Sequence
8	1303.6	51.3	1755	6 AK001790	AK001790 Homo sapi
9	1270.6	50.0	2322	6 AK398774	AK398774 Sequence
10	1224.2	48.2	2064	6 AK000522	AK000522 Homo sapi
11	1216.4	47.9	2061	6 AK405828	AK405828 Sequence
12	1028	40.5	1416	6 AK451362	AK451362 Sequence
13	1012	39.8	160903	2 AP002776	AP002776 Homo sapi
14	1012	39.8	171566	2 AP001885	AP001885 Homo sapi
15	1003.4	39.5	1416	9 AB072360	AB072360 Homo sapi
16	797.6	31.4	2736	10 BC028922	BC028922 Mus muscu
17	779.6	30.7	1026	6 AX086034	AX086034 Sequence
18	573.6	22.6	580	9 HUMY153C10	HUMY153C10 Homo sapi
19	558.6	22.0	599	6 AX398772	AX398772 Sequence
20	397	15.6	409	6 AX398771	AX398771 Sequence
21	336	13.2	426	6 AX398765	AX398765 Sequence
22	324.4	12.8	334	6 AX337360	AX337360 Sequence
23	241	9.5	494	6 AX398773	AX398773 Sequence
24	234.2	9.2	279	11 G22634	G22634 human SRS W
25	216.6	8.5	4417	3 AB036834	AB036834 Drosophila
26	212	8.3	1949	6 AX223960	AX223960 Sequence
27	212	8.3	2280	9 AB072356	AB072356 Homo sapi
28	212	8.3	2433	9 AK095421	AK095421 Homo sapi
29	212	8.3	3488	6 AX406972	AX406972 Sequence
30	212	8.3	3817	9 AB072355	AB072355 Homo sapi
31	212	8.3	6374	6 AX180874	AX180874 Sequence
32	208.4	8.2	1052	6 AX223966	AX223966 Sequence
33	208.4	8.2	1711	6 AX223964	AX223964 Sequence
34	201.8	7.9	1796	9 AB072358	AB072358 Homo sapi
35	182.4	7.2	2260	6 AX180876	AX180876 Sequence
36	174.4	6.9	1771	6 AX202239	AX202239 Sequence
37	142.8	5.6	386	9 AF484838	AF484838 Homo sapi
38	142	5.6	113474	2 AC013932	AC013932 Drosophila
39	142	5.6	181132	3 AC008206	AC008206 Drosophila
40	142	5.6	227398	3 AE003750	AE003750 Drosophila
41	121.4	4.8	5463	9 AB037719	AB037719 Homo sapi
42	104.4	4.1	1390	6 AX237020	AX237020 Sequence
43	103.4	4.1	531	6 AX237022	AX237022 Sequence
44	98.8	3.9	1236	6 AX354534	AX354534 Sequence
45	98.8	3.9	1693	9 BC007771	BC007771 Homo sapi

ALIGNMENTS

```
RESULT 1
LOCUS AK074432 2808 bp mRNA linear PRI 15-FEB-2002
DEFINITION Homo sapiens cDNA FLJ23852 fls, clone KAT12021.
ACCESSION AK074432
VERSION AK074432.1 GI:18677033
KEYWORDS Oligo capping; fls (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to
mrna, clone_11b:KAT clone:KAT12021.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Watanabe,K., Kunagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
```

TITLE Nakamura,Y., Isogai,T. and Sugano,S.
JOURNAL MEDO human cDNA sequencing project
REFERENCE 2 (bases 1 to 2808)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Ohayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
JOURNAL Direct Submission
Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
(E-mail:cdna@iims.u-tokyo.ac.jp; Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT MEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
Source
1..2808
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KATO12021"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_id="KAT"
/note="cloning vector: pME18SFL3"
97..2076
/note="unnamed protein product"
/codon_start=1
/protein_id="BA85080.1"
/db_xref="GI:18677034"
/translation="MALIVTSRSPSPGSGASTPYGPMDOAVORRSRLORROFAVLNGA
VLIOLOGDNDAAEASEPTEKAPSEELHGDQTFQSGQSPQKOEORHLHMY
OLLRPODDILAEOLLEAPRPRLYLIVSTRGEGLSODETVLLIVDPDSSPSC
LGLVPLMSDTQVYLDGDFSVTSQSGSIFKPISTQIMMATLYLHOACAEALSG
LYPGSALTNASHYGERLNSQSLMBWIMADLESRLPESAGSSSEDOEMOKIR
AELMKVLDVDSLETSVKELRLQALELRLGLPQOYDFIDNOMLLVAORDRSRIEP
HLTGSEMANANLEILORNVRTHILNMRKIDNFEPRFTYHNVRIMDESADLLPM
KETHRFLEAARQGTIVLVKMGVSRSATVLAAMKQGESLEBOALRHVDELRFA
RPNPGEFROLOITQIGILTASROSHWQKVGAGSPENHAPVSTPEPPLPEPEGG
EEKVGMESQAPKREPGRPRINLRGVWRSIILEPSTLESTSETDMPVESH
ESSHEPPLQEPOLARTKGGQYDRGPRALKRSQSVYTLQSAVAVNRTQAFQEDQ
GGQGGGEPCTISIPRSKRVKQASVHDSSEGEA"
BASE COUNT 585 a 912 c 823 g 488 t
ORIGIN
Query Match 84.7%; Score 2151; DB 9; Length 2808;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 2506; Conservative 0; Mismatches 5; Indels 292; Gaps 6;

Dh 310 GAGGAGCTCCACGGGGACCAAGACAGACTTCTGGGCAAGATCCAGATCCCAAGAGCAG 369
Qy 361 GAGGAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 420
Dh 370 GAGGAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 429
Qy 421 CGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 480
Dh 430 CGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 489
Qy 481 TCTACACGAG 540
Dh 490 TCTACACGAG 549
Qy 541 CTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600
Dh 550 CTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 609
Qy 601 CAGGTGACTTATGATGAG 660
Dh 610 CAGGTGACTTATGATGAG 669
Qy 661 TTTGACGCCATCTCCATCCAGACCATGT----- 688
Dh 670 TTTGACGCCATCTCCATCCAGACCATGTGGCCACATCTGATTTGACCAAGCATGT 729
Qy 689 ----- 688
Dh 730 GAGGAGCAGCTACGACAGCGGCTTTGACCGGGGTGGAGTGCCTCAGCTGGGCGAGCCAC 789
Qy 689 ----- 688
Dh 790 TACACAGAGAGACTGAACCTCCAAACAGACAGCTGCCTCAATAGTGAAGAGAGCTATGGCCGAC 849
Qy 689 -----GGTCCCTCAAGAGAGAGAGAGATG 711
Dh 850 CTGAGACTCTCTGGGCTCCACAGCGCCAGAGCTGCGGGGTCTCTCAAGAGAGAGAGATG 909
Qy 712 GACACAGCAGCTCGTCTGAGCTGAGTGAAGAGTGGTGAAGTGAAGTGAAGTGAAGTGAAG 771
Dh 910 GACACAGCAGCTCGTCTGAGCTGAGTGAAGTGGTGAAGTGAAGTGAAGTGAAGTGAAG 969
Qy 772 ACTTCCAAAGATTCGCGCAGAGCTGTGAGAGTGGCTGGGCTCCCTCCAGCAGTAC 831
Dh 970 ACTTCCAAAGATTCGCGCAGAGCTGTGAGAGTGGCTGGGCTCCCTCCAGCAGTAC 1029
Qy 832 CGTGACTCTATGACAAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
Dh 1030 CGTGACTCTATGACAAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
Qy 892 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAAGAGCAGCAAACTGGAGAGAGTGCAG 951
Dh 1090 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAAGAGCAGCAAACTGGAGAGAGTGCAG 1149
Qy 952 AGGAGAGGCTCACCAATCTTGAACATGCGCCGAGAGATTGACAACTTACCTCTGAG 1011
Dh 1150 AGGAGAGGCTCACCAATCTTGAACATGCGCCGAGAGATTGACAACTTACCTCTGAG 1209
Qy 1012 CGCTTACCTTACCAATGTCGCTGCTGAGAGAGAGATGCGCCAGCTGCTGCTGCTGCTGCT 1071
Dh 1210 CGCTTACCTTACCAATGTCGCTGCTGAGAGAGATGCGCCAGCTGCTGCTGCTGCTGCTGCT 1269
Qy 1072 TGGAGAGAGAGCAGCAGCTTATGAGAGTGCAGAGAGCAGAGAGCAGCAGCTGCTGCTGCTGCT 1131
Dh 1270 TGGAGAGAGAGCAGCAGCTTATGAGAGTGCAGAGAGCAGAGAGCAGCAGCTGCTGCTGCTGCT 1329
Qy 1132 CACTGCAAGATGGGCTCAGCGCTCAGCGGCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1191
Dh 1330 CACTGCAAGATGGGCTCAGCGCTCAGCGGCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1389
Qy 1192 TACGAATGACAGCTGAG 1251
Dh 1390 TACGAATGACAGCTGAG 1449

```

OY 1252 CCCAACCCCTGGCTTCGCGCCAGCTGAGATCTACAGGGGATCTTGAGC ----- 1302
Db 1450 CCCAACCCCTGGCTTCGCGCCAGCTGAGATCTACAGGGGATCTTGAGCAGCCGCC 1509
OY 1303 ----- 1302
Db 1510 CAGAGCCATGCTGGAGAGAGAAAGGGTGTCTCCAGAGGAGACCCAGCCCTT 1569
OY 1303 -----GCCAGAACCTGAGGGTGGTGGGGAGAGAG 1333
Db 1570 GAACTCTACACCATTCCTCCACCTTCCTCCGACAGACCTGGGGGTGGGGAGAGAG 1639
OY 1334 GTTGTAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
Db 1630 GTTGTAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1667
OY 1394 GTATTAACCTTCGAGGGGTGATGAGGTCCATCAGTCTTGTGAGAGCCCTCTTGGAGCTG 1453
Db 1688 GTATTAACCTTCGAGGGGTGATGAGGTCCATCAGTCTTGTGAGAGCCCTCTTGGAGCTG 1746
OY 1454 GAGAGACCTTCAGTACAGTACATGATGAGAGAGTCTTCTTCCACAGAGTCTTACA 1513
Db 1747 GAGAGACCTTCAGTACAGTACATGATGAGAGAGTCTTCTTCCACAGAGTCTTACA 1805
OY 1514 TGAAGAGCTTCGAGAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAGAG 1573
Db 1806 TGAAGAGCTTCGAGAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAG 1865
OY 1574 CAGGGGGCTTCAGAGCCCTTCCAGAGTCCAGTCCAGTCCAGTCCAGAGAGAGTGC 1633
Db 1866 CAGGGGGCTTCAGAGCCCTTCCAGAGTCCAGTCCAGTCCAGTCCAGAGAGAGTGC 1925
OY 1634 CGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
Db 1926 CGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1985
OY 1694 GGGAGAGCCCTTCAGTTCCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753
Db 1986 GGGAGAGCCCTTCAGTTCCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2045
OY 1754 GCATGACAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1813
Db 2046 GCATGACAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2105
OY 1814 ACTGAAGAGATTCACAACTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1873
Db 2106 ACTGAAGAGATTCACAACTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2165
OY 1874 CTCTACAGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1933
Db 2166 CTCTACAGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225
OY 1934 CTACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1993
Db 2226 CTACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2285
OY 1994 GTCTGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2053
Db 2286 GTCTGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2345
OY 2054 CTTTTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2113
Db 2346 CTTTTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2405
OY 2114 GGCAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2173
Db 2406 GGCAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2465
OY 2174 CTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2233
Db 2466 CTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2525
```

```

OY 2234 GTGGCCCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2293
Db 2526 GTGGCCCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2585
OY 2294 CAGATCGCTTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2353
Db 2586 CAGATCGCTTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2645
OY 2354 CTGTGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2413
Db 2646 CTGTGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2705
OY 2414 CTGGCTTCGCTAGTACAGTTCATGAGTTCATGAGTTCATGAGTTCATGAGTTCAT 2473
Db 2706 CTGGCTTCGCTAGTACAGTTCATGAGTTCATGAGTTCATGAGTTCATGAGTTCAT 2765
OY 2474 AAACACATTTTCATATTTGTTAAAAA 2516
Db 2766 AAACACATTTTCATATTTGTTAAAAA 2808
```

RESULT 2

```

AX099939
LOCUS AX099939
DEFINITION Sequence 21 from Patent WO012004.
ACCESSION AX099939
VERSION AX099939.1 GI:13538949
KEYWORDS
SOURCE
ORGANISM human.
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2781)
Yue, H., Tang, Y.-T., Bandman, O., Hillman, J.L., Baughn, M.R., Azimzal, Y. and Lu, D.A.

TITLE Protein phosphatase and kinase proteins
JOURNAL Patent: WO 012004-A 21 22-MAR-2001;
Incyte Genomics, Inc. (US)

FEATURES
source 1..2781
Location/Qualifiers

BASE COUNT 576 a 906 c 820 g 479 t
ORIGIN
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID NO: 5039718CB1"

Query Match 83.3%; Score 2115; DB 6; Length 2781;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 5; Indels 292; Gaps 6;

```

OY 10 GCGGGTCCAGACTGTCCCGGGGGTTGAGGAGAGGGGCGTGGCCGAGCCAGG 69
Db 10 GCGGGTCCAGAGACTGT-CCGCGGGGTTGAGGAGAGGGGCGTGGCCGAGCCAGG 68
OY 70 TGCTGCGGGCTGGCTGCATAGGCTTGCTACAGTGAAGCCGTTGCCCGGAGCGAGC 129
Db 70 TGCTGCGGGCTGGCTGCATAGGCTTGCTACAGTGAAGCCGTTGCCCGGAGCGAGC 128
OY 69 TGCTGCGGGCTGGCTGCATAGGCTTGCTACAGTGAAGCCGTTGCCCGGAGCGAGC 128
Db 69 TGCTGCGGGCTGGCTGCATAGGCTTGCTACAGTGAAGCCGTTGCCCGGAGCGAGC 128
OY 130 GCCTTCACAGCCGCTGGGGGCTGGAGACAGGAGGCTCCAGGAGAGAGTGAAGTCAAGCA 189
Db 129 GCCTTCACAGCCGCTGGGGGCTGGAGACAGGAGGCTCCAGGAGAGAGTGAAGTCAAGCA 188
OY 190 AGGAGAGAGTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGG 249
Db 189 AGGAGAGAGTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGG 248
OY 250 GATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
Db 249 GATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 308
OY 310 CAGGGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
Db 309 CAGGGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
```


Oy	370	AGGAGCAGCTCGACACTCATGTCGTACGACTGCTGAGAGCCGACAGATGACATCCGCTTGCA	429
Db	369	AGGGAGAGCACTGGACACCTCAATGATGACGTCTGAGAGCCGACAGATGACATCCGCTTGCA	428
Oy	430	GCCAGCTGGAGAGACCCCGGCTCCCGGCTCCGCTACCTCTGTAAGTTTCTACAGCA	489
Db	429	GCCAGCTGGAGAGCACCCCGGCTCCCGGCTCCGCTACCTCTGTAAGTTTCTACAGCA	488
Oy	490	GAAGAGAGAAGTCTGAACCCAGATGAGACGGTCTCTGGGCGTGGATTTCCCTGACAGC	549
Db	489	GAAGAGAGAAGTCTGAACCCAGATGAGAGGGTCTCTGGGCGTGGATTTCCCTGACAGC	548
Oy	550	AGCTGCCCAAGCTGCACCCCGGGGCGTGGTCTTGCCCTCTGGAGTGAACACCAGAGTGC	609
Db	549	AGCTGCCCAAGCTGCACCCCGGGGCGTGGTCTTGCCCTCTGGAGTGAACACCAGAGTGC	608
Oy	610	TTAGATGAGAGCGGGGGCTTCACCTGAGAGTCTGTGGGGCAAGCCGGATCTTCAAGCCC	669
Db	609	TTAGATGAGAGCGGGGGCTTCACCTGAGAGTCTGTGGGGCAAGCCGGATCTTCAAGCCC	668
Oy	670	ATTCCTCATCCAGACATGT-----	688
Db	669	ATTCCTCATCCAGACATGTGGGCGACACTGCAGGTAATTGCACCAAGATGTGAGGCACT	728
Oy	689	-----	688
Db	729	CTAGAGCAGCGGCTTCTTACCGGGGTGGCAGTGGCCCTCACCCTGGGCCACGCACTACCAAGAG	788
Oy	689	-----	688
Db	789	AGACTGAACCTCCGAAACAGAGACTGCTCAATGATGAGACGGCTATMGCCGACCTGAGTCT	848
Oy	689	-----GTTCTCGAAGACAGACAGATGGAGAGGCG	720
Db	849	CTGCGGCTCCCACTGCGCCAGCCTGGCGGGTCTCTCGAAGACAGACAGATGGAGAGGCG	908
Oy	721	ATTCGTCGTAGCTGTGGAAAGTGTGGATGTGCATGACACTTGGAGAGTGTCACTTCCAA	780
Db	909	ATTCGTCGTAGCTGTGGAAAGTGTGGATGTGCATGACACTTGGAGAGTGTCACTTCCAA	968
Oy	781	GAGATCCGCGCAGGCTGTGGAGCGTGGGCGTGGGGCTCCCTCCAGACATACCGTGACTTC	840
Db	969	GAGATCCGCGCAGGCTGTGGAGCGTGGGCGTGGGGCTCCCTCCAGACATACCGTGACTTC	1028
Oy	841	ATGCACACAGAGATGCTGCTGTGTGTGGACAGCGGGACCGAGCTCCCGCATCTTCCCC	900
Db	1029	ATGCACACAGAGATGCTGCTGTGTGTGGACAGCGGGACCGAGCTCCCGCATCTTCCCC	1088
Oy	901	CACCTTACCTGGGCTTAAGATGTGAACCGACCAACTTGTGAAGAGCTGCAGAGAACAGG	960
Db	1089	CACCTTACCTGGGCTTAAGATGTGAAGCGACCAACTTGTGAAGAGCTGCAGAGAACAGG	1148
Oy	961	GTCACCCACATCTTGACATGTGGCCGGGAGATTTGACACTTCTACCTGAGGCGTTTCAAC	1020
Db	1149	GTCACCCACATCTTGACATGTGGCCGGGAGATTGACAACTTCTACCTGAGGCGTTTCAAC	1208
Oy	1021	TACCACAATGTGGCTCTGTGGATGAGAGATGGGCCACAGTCTGCCCACTGGAAGAG	1080
Db	1209	TACCACAATGTGGCTCTGTGGATGTAGAGATGGGCCCAAGCTCTGCCCACTGGAAGAG	1268
Oy	1081	ACGACCGCTTATGTAGGCGTGAAGAGCAACAGGACCAAGTGTGTGTGCACATGCAAG	1140
Db	1269	ACGACCGCTTATGTAGGCGTGAAGAGCAACAGGACCAAGTGTGTGTGCACATGCAAG	1328
Oy	1141	ATGGAGCGTACAGCCGCTACAGGGCCACAGTGTGGCTATATGCATTAAGCAAGTAAATGC	1200
Db	1329	ATGGAGCGTACAGCCGCTACAGGGCCACAGTGTGGCTATATGCATTAAGCAAGTAAATGC	1388
Oy	1201	AGCTGAGAGCAGGCGCTGCGCCACAGTGCAGAGACTCGGCGCCATCGCCCGCCAACTT	1260
Db	1389	AGCTGAGAGCAGGCGCTGCGCCACAGTGCAGAGACTTCGCGCCCATGCGCCCGCCAACTT	1448

QY	1261	GGTTTCCTCGCGCCAGCTGCGAGATCTACAGGGCATCTCTACG-----	1302
Db	1449	GGCTTCCTCGCGCCAGCTGCGAGATCTACAGGGCATCTCTACGCGCCAGCGACCAT	1508
QY	1303	-----	1302
Db	1509	GTCTGGAGCAGAAAGTGGGTGGGGTCTCCCCAGAGAGACACCAGCCCCCTGAACTCTCT	1568
QY	1303	-----GCCAAGAACCTGAGGGGTGTGGGGAGAGAGAAAGTTGTAGGC	1342
Db	1569	AACACATTCCCACTCTTTCGGCCAGAAACCTGAGGGGTGTGGGGAGAGAAAGTTGTAGGC	1628
QY	1343	ATGGAAGAGAGCCAGCGCAGACCCCGAAAGAAAGAGCTTGGGGCCACGGGGCAGATATAACC	1402
Db	1629	ATGGAAGAGAGCCAGCGCAGACCCCGAAAGAAAGAGACCTCTGGG--CCAGCGGCCACTATATAACC	1686
QY	1403	TCCGAGGGGTATAGGTTCCATCATCACTCTTTGGAGGCCCTCTTGGAGAGCTGGAGAGCACC	1462
Db	1687	TCCGAGGGGTATAGGTTCCATCATCACTCTTTGGAGGCCCTCTT--GGAGCTGGAGAGCACC	1745
QY	1463	TCAGTAGACACGATACATATGCCAGAGGGTCTTCTTCCACGAGTCTTGACATGAGAAGCC	1522
Db	1746	TCAG--AGACGAGTACATATGCCAGAGGGTCTTCTTCCACGAGTCTTGACATGAGAAGCC	1804
QY	1523	TCTGACAGCCCTTCCACACGCTTGCAAGAGCAAGAGAGCGCAGCAGGTGAGAGAGGGCC	1582
Db	1805	TCTGACAGCCCTTCCACACGCTTGCAAGAGCAAGAGAGCGCAGCAGGTGAGAGAGGGGCC	1864
QY	1583	TCAGCTCTCCCTGAAAGTCCCGCCAGTACAGTGGTTACCTCCAGGGCAGTCCGTGTGGG	1642
Db	1865	TCAGCTCTCCCTGAAAGTCCCGCCAGTACAGTGGTTACCTCCAGGGCAGTCCGTGTGTGGC	1924
QY	1643	CAACCGGACCCAGGCGCTTCCAGAGAGAGAGAGAGGGGCGAGGGGCGAGGAGAGACC	1702
Db	1925	CAACCGGACCCAGGCGCTTCCAGAGAGAGAGAGAGGGGCGAGGGGCGAGGAGAGACC	1984
QY	1703	CTGCAATTCTCTACAGCCCGCAGGTTCCGGAAAGGTGTGTAGACAGCCAGCCAGTGCATGAC	1762
Db	1985	CTGCAATTCTCTACAGCCCGCAGGTTCCGGAAAGGTGTGTAGACAGCCAGGCTGCATGACAG	2044
QY	1763	TGGAGAGAGAGGGGCGAGGCGTGAAGCTGCACATGCGCCAGCGCTCCAGTACCTGAACAG	1822
Db	2045	TGGAGAGAGAGGGGCGAGGCGTGAAGCTGCACATGCGCCAGCGCTCCAGTACCTGAACAG	2104
QY	1823	GATCCACACATCTCTTGGAGAGAAACCCCTCACTGCTTTGCCGACACATTCCTCTCACT	1882
Db	2105	GATCCACACATCTCTTGGAGAGAAACCCCTCACTGCTTTGCCGACACATTCCTCTCACT	2184
QY	1883	CCGCCCCATACCCGCTCACTACAGCCTCACTCCACCCCTGTACTAGAGGCTCACTCC	1942
Db	2165	CCGCCCCATACCCGCTCACTACAGCCTCACTCCACCCCTGTACTAGAGGCTCACTCC	2224
QY	1943	CACCCCTGTCACTACAGCCTCACTCCCTAAGGCTTAATGCCAAGGCCATGTGCGCTG	2002
Db	2225	CACCCCTGTCACTACAGCCTCACTCCCTAAGGCTTAATGTCCAGAGCCCATGTCTGGCTG	2284
QY	2003	TCCAAGGGCTCAAACTTTTAACTGGGATGTGTAGAGAGGACTAAAGTATACCTTTGGGG	2062
Db	2285	TCCAAGGGCTCAAACTTTTAACTGGGATGTGTAGAGAGGACTAAAGTATACCTTTGGGG	2344
QY	2063	GCAACAGCACCTTAGTTTCAATTCCTCAACTCTAGCCCTGCACACTACCTGTGGCAGGAA	2122
Db	2345	GCAACAGCACCTTAGTTTCAATTCCTCAACTCTAGCCCTGCACACTACCTGTGGCAGGAA	2404
QY	2123	TGAAAACAGAGACTTCGGGTGCAAAAAGGGTCAAGGCTCCGACCCCGGCCCTCTCCCTGCA	2182
Db	2405	TGAAAACAGAGACTTCGGGTGCAAAAAGGGTCAAGGCTCCGACCCCGGCCCTCTCCCTGCA	2464
QY	2183	CCTCTCTGCTCTCCGAGTTTCAATTCCTGGAAGACAGCCAGGCCAGCAACAGATGGCCCC	2242
Db	2465	CCTCTCTGCTCTCTCCGAGTTTCAATTCCTGGAAGACAGCCAGGCCAGCAACAGATGGCCCC	2524
QY	2243	AAAGCAGGAGGATTCCTCAGGGCCCAAGCCGCGGAGGCTGGAAAGGCTGGCAGATGCT	2302

|||||
Db 2525 AAGGCGAGGAGATCCTCAGGCGCCGAGCGGAGGCTGGAAGGCTGCGACATCGCT 2584
Qy 2303 TCCCTCATCCACCCGACCGGCTCCAGCTCTTGTGCTGCCGCCAGACCTCTGTGACAC 2352
Db 2585 TCCCTCATCCACCCGACCGGCTCCAGCTCTTGTGCTGCCGCCAGACCTCTGTGACAC 2644
Qy 2363 CAGCCGAGATCACAGGCGACGAGATAGTCTTCTTGTCTCTTCTGCTGCTGCT 2422
Db 2645 CAGCCGAGATCACAGGCGACGAGATAGTCTTCTTGTCTCTTCTGCTGCTGCT 2704
Qy 2423 GCGTACGATTTTTCATACGCTTACAGATCTGCTGCTTGTCTGAGAATAAACAAT 2482
Db 2705 GCGTACGATTTTTCATACGCTTACAGATCTGCTGCTTGTCTGAGAATAAACAAT 2764
Qy 2483 TTTTCATA 2489
Db 2765 TTTTCATA 2771
RESULT 3
LOCUS AK094226 2905 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ36907 fls, clone BRACE2003800, weakly similar
to MAP kinase phosphatase.
ACCESSION AK094226
VERSION AK094226.1 GI:21753246
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_11b:BRACE2
clone:BRACE2003800.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiraio,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
TITLE NEDO human cDNA sequencing project
JOURNAL
REFERENCE 2 (bases 1 to 2905)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL
COMMENT Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team): 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB): cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1. .2905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2003800"
/tissue.type="cerebellum"
/clone_11b="BRACE2"
/note="cloning vector: pME18SFL3"
646..2187
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04314.1"
/db_xref="GI:21753247"
/translation="MAPKIPASSPSCCTGLVPLPMSDPTQVYLDGGGSEVTSGGGSR
IFPISIQTMATLVQLHQACEALISGLVPGSALTWASHYQERLNSBOSCLNEMTA

MADLESLRPPSAPGSSSEQEQEQA1RAELMVLVDLSVTSKEIRQALRLGL
PLOQYRDEIDNOMLLVARDRASRIEPLHYLTSENNAALELQNRKTHILNMAE
IDNPERFETHVRLWDESSAOLLPWKETHRFIFAAKROGTHVHCKMGVSRSA
TVLAVAMKOYECLEQALRHVQELRLAPRNPFLQLOLYQGLITLARSRSYWEKV
GGVSEBHPAPEVSTPPPLPEPEGGKVVGMESCAAPKEEPFRIRLQGM
RSLSLEPSELESTSETSDMPVFSSHESHEPRLQFQPLARKTGQGVDPDPA
LKSROSVYVLQGSVVAVNARQAFQEOEGOGGOGCPICISSTPRFKVVRQASVHDS
EEGEA"
BASE COUNT 591 a 940 c 867 g 507 t
ORIGIN
Query Match 76.0%; Score 1930.4; DB 9; Length 2905;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2457; Conservative 0; Mismatches 6; Indels 443; Caps 8;
Qy 27 CCGGGGGTTAGGGAAGGGGCGTGCCTGCGACCCAGGTGCTCGCGGCTGCTC 86
Db 3 CCGGGGGTTAGGGAAGGGGCGTGCCTGCGACCCAGGTGCTCGCGGCTGCTC 62
Qy 87 CATGGCCCTGCTACAGTGAAGCCGTTGCCCCGGGGAGCGGCGCTCCACGCGTGGG 146
Db 63 CATGGCCCTGCTACAGTGAAGCCGTTGCCCCGGGGAGCGGCGCTCCACGCGTGGG 122
Qy 147 GCCCTG----- 152
Db 123 GCCCTGGAATTCGTAGAGAGGAGGAGGAGCGCCCTCCCGCCCTACCCCTGCCCTGG 182
Qy 153 -----GACACAGCGGCGTCCAGCGAAGAGTGAAGTCCAGCAGAGAGAG 197
Db 183 GCTCTCTCTCGCAGAGACACGAGGCGTCCAGCGAAGAGTGAAGTCCAGCAGCA 238
Qy 198 CTTGCGGTGCTCGTGGGCGTCTGCGACTGTCAGAGTGAAGGGGAGCAATGATGTC 257
Db 239 TTTTGGGTGCTCGTGGGCGTCTGCGACTGTCAGAGTGAAGGGGAGCAATGATGTC 298
Qy 258 AGCAGAGCCAGTTCTGAGCCACAGAGAAAGCCCGAGTGAAGAGAGTCCACGGGGA 317
Db 299 AGCAGAGCCAGTTCTGAGCCACAGAGAAAGCCCGAGTGAAGAGAGTCCACGGGGA 358
Qy 318 CCAACAGACTTGGGCGAAGATCCCGAGTCCCGACAGAGAGAGAGAGCAGAGCAGCA 377
Db 359 CCAACAGACTTGGGCGAAGATCCCGAGTCCCGACAGAGAGAGAGAGCAGAGCAGCA 418
Qy 378 CCTGACCTCATGTACAGCTGAGGCGCCAGAGATGATCCGCTGCGAGCCAGCT 437
Db 419 CCTGACCTCATGTACAGCTGAGGCGCCAGAGATGATCCGCTGCGAGCCAGCT 478
Qy 438 GGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTTCTACAGAGAAGAGA 497
Db 479 GGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTTCTACAGAGAAGAGA 538
Qy 498 AGGTCTAGCAGAGATGAGAGGCTCTCTGCGCTGATTTCCCTGACAG----- 548
Db 539 AGGTCTAGCAGAGATGAGAGGCTCTCTGCGCTGATTTCCCTGACAGAGTTGCA 598
Qy 549 ----- 548
Db 599 GCAGGAGAGGAAGAGGGGCAAGAGGTGAGGAGAGGGGAGCAATGCTTCCCTC 658
Qy 549 -----CAGTCCCGCAGCTGCAAGCTGAGCTGCTTCCCTGAGAGTGA 598
Db 659 TGTCCCTGACAGCTCCCGCAGCTGCAAGCTGAGCTGCTTCCCTGAGAGTGA 718
Qy 599 CCCAGGTGATTTAGATGAGAGCGGGGCTTACAGCTGAGCTGCTGAGGAGAAAGCCGGA 658
Db 719 CCCAGGTGATTTAGATGAGAGCGGGGCTTACAGCTGAGCTGCTGAGGAGAAAGCCGGA 778
Qy 659 TCTTCAAGCCATCTCATCCAGACATGTGGGACAGCTCAGAGTATGTGACCAAGAT 688
Db 779 TCTTCAAGCCATCTCATCCAGACATGTGGGACAGCTCAGAGTATGTGACCAAGAT 838
Qy 689 ----- 688

Dh 839 GTGAGGAGCTTAGGAGGAGGCTTGTACGGGTGGCAGTGGCCCTACCTGGGCCAGCC 898
Qy 689 ----- 688
Dh 899 ACTACAGAGAGAGACTGAAGCTCGAAGCAGAGCTGCCAATGAGTGGACGGCTATGGCCG 958
Qy 689 -----GTCCTCGACAGAGAGAGA 709
Dh 959 ACCTGAGTCTCTGCGGCTCCGACGGCCGAGCCTGGCGGGTCTCTGAGACAGAGAGAGA 1018
Qy 710 TGGAGAGAGGAGATCCGCTGCTGAGCTGTGGAAGTGTGGATGTGATGAGTGCAGAGAGAGT 769
Dh 1019 TGGAGAGAGGAGATCCGCTGCTGAGCTGTGGAAGTGTGGATGTGATGAGTGCAGAGAGT 1078
Qy 770 TCATTTCGAAGAGATCCGCTGCTGAGCTGTGAGTGGCCGAGGGGCTCCCTCCACAGT 829
Dh 1079 TCATTTCGAAGAGATCCGCTGCTGAGCTGTGAGTGGCCGAGGGGCTCCCTCCACAGT 1138
Qy 830 ACCGTGACTTCATGCAACAAACAGATGCTGCTGTGGTGGCACAAGCGGAGACGAGCTGCC 889
Dh 1139 ACCGTGACTTCATGCAACAAACAGATGCTGCTGTGGTGGCACAAGCGGAGACGAGCTGCC 1198
Qy 890 GCATCTTCCCGCCTCTACCTGCTGGCTGAGTGGAGAACGAGAACGAGTGCAGAGAGTGC 949
Dh 1199 GCATCTTCCCGCCTCTACCTGCTGGCTGAGTGGAGAACGAGAACGAGTGCAGAGAGTGC 1258
Qy 950 AGAGAGACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTGACACTTCTACCTG 1009
Dh 1259 AGAGAGACAGGGGTACCCACATCTTGAACATGAGCCCGGAGATTGACACTTCTACCTG 1318
Qy 1010 AGCGCTTCACCTACCAACATGTGGGCTGTGGGATGAGAGTGGCCAGCTGTGCTGCGC 1069
Dh 1319 AGCGCTTCACCTACCAACATGTGGGCTGTGGGATGAGAGTGGCCAGCTGTGCTGCGC 1378
Qy 1070 ACTGGAAGAGAGACGACCGCTTCATTGAGGCTGCAAGAGACAGGAGACCGAGTGTGG 1129
Dh 1379 ACTGGAAGAGAGACGACCGCTTCATTGAGGCTGCAAGAGACAGGAGACCGAGTGTGG 1438
Qy 1130 TCCAGTGGCAAGTGGGGCTGACAGCCGCTAGCGGCCACAGTGTGCTTATGCTCATGAAC 1189
Dh 1439 TCCAGTGGCAAGTGGGGCTGACAGCCGCTAGCGGCCACAGTGTGCTTATGCTCATGAAC 1498
Qy 1190 AGTAGAATGACGCTGAGAGAGGCTGCGCCAGCTGAGAGAGTCCGGGCCATCGGCC 1249
Dh 1499 AGTAGAATGACGCTGAGAGAGGCTGCGCCAGCTGAGAGAGTCCGGGCCATCGGCC 1558
Qy 1250 GCGCCAAACCTTGGCTTCTGCGGCCAGCTGACAGATCTACAGGGCATCTCTGAAGC 1302
Dh 1559 GCGCCAAACCTTGGCTTCTGCGGCCAGCTGACAGATCTACAGGGCATCTCTGAAGC 1618
Qy 1303 ----- 1302
Dh 1619 GCCAGAGCATGTCTGGAGAGCAAGAAATGGGTGGGGTCTCCCGAGAGAGACCCAGCC 1678
Qy 1303 -----GCCAGAACTGAGGGTGTGGGAGAGAGA 1331
Dh 1679 CTGAAGTCTTACACCATTTCCACACTTCTCCGCAAGAACTGAGGGTGTGGGAGAGAGA 1738
Qy 1332 AGGTTGATGAGATGGAAGAGAGAGCCAGGACGCCGAAAGAAAGAGCTTGGGGCCACGGGGC 1391
Dh 1739 AGGTTGATGAGATGGAAGAGAGAGCCAGGACGCCGAAAGAAAGAGCTTGGGGC 1796
Qy 1392 ACATATAAACCTCCGAGGGGTCATGAGGTTCATGAGTCTTCTGAGGCCCTCTTGGAGAGC 1451
Dh 1797 ACATATAAACCTCCGAGGGGTCATGAGGTTCATGAGTCTTCTGAGGCCCTCTTGGAGAGC 1855
Qy 1452 TGGAGAGACCTCATGATAGACAGTGCAGAGAGTCTTCTTCCAGAGTCTTCA 1511
Dh 1856 TGGAGAGACCTCATGATAGACAGTGCAGAGAGTCTTCTTCCAGAGTCTTCA 1914
Qy 1512 CATGAAGAGGCTTGGAGAGGCTTCCAGCTTGCAGAGAGCAAGGAGGCGCAGAGGTG 1571
Dh 1915 CATGAAGAGGCTTGGAGAGGCTTCCAGCTTGCAGAGAGCAAGGAGGCGCAGAGGTG 1974

Qy 1572 GACAGGGGGCTCAGCCTGCTGAAGTCCCGCCAGTCAAGTGGTTACCTCCAGGGCAGT 1631
Dh 1975 GACAGGGGGCTCAGCCTGCTGAAGTCCCGCCAGTCAAGTGGTTACCTCCAGGGCAGT 2034
Qy 1632 GCCGTGGTGGCCAAACCGGAGCCGCTTCCAGAGACAGAGAGAGGAGGAGGAGG 1691
Dh 2035 GCCGTGGTGGCCAAACCGGAGCCGCTTCCAGAGACAGAGAGAGGAGGAGGAGG 2094
Qy 1692 CAGGAGAGCCCTGCATTTCTCTACGCGCCAGGTCCGGAAGTGTGAGAGAGCCAGC 1751
Dh 2095 CAGGAGAGCCCTGCATTTCTCTACGCGCCAGGTCCGGAAGTGTGAGAGAGCCAGC 2154
Qy 1752 GTGCATGACAGTGGAGAGAGGAGGAGGCTGAGGCTTACACATGAGCCAGCTCCCTG 1811
Dh 2155 GTGCATGACAGTGGAGAGAGGAGGAGGCTGAGGCTTACACATGAGCCAGCTCCCTG 2214
Qy 1812 ACAGTGAAGAGGATCCACACTCTTGGAGAAACACCTCAGCTGTGTTGCCGACAGAT 1871
Dh 2215 ACAGTGAAGAGGATCCACACTCTTGGAGAAACACCTCAGCTGTGTTGCCGACAGAT 2274
Qy 1872 TCCCTCAGCTCCGCGCCATACCGCTCACTACAGCTCAGCTCCAGCCCTGTCTACG 1931
Dh 2275 TCCCTCAGCTCCGCGCCATACCGCTCACTACAGCTCAGCTCCAGCCCTGTCTACG 2334
Qy 1932 GCCTCACTCCGACCCCTGTCACTACAGCTCAGCTCCTTACAGCCTTAAGTCCAGGCCC 1991
Dh 2335 GCCTCACTCCGACCCCTGTCACTACAGCTCAGCTCCTTACAGCCTTAAGTCCAGGCCC 2394
Qy 1992 AATGTGCTGTGCAAGGGCTCAAGACTTCTTAATGAGGATGGTGAAGGAGCTGAAG 2051
Dh 2395 AATGTGCTGTGCAAGGGCTCAAGACTTCTTAATGAGGATGGTGAAGGAGCTGAAG 2454
Qy 2052 TACCTTTGGGGCAACAGACCTTATGTTTCACTCTCACTCTAGCCCTGACACTCACT 2111
Dh 2455 TACCTTTGGGGCAACAGACCTTATGTTTCACTCTCACTCTAGCCCTGACACTCACT 2514
Qy 2112 GTGGACAGGATGAAGAAACAGAGCTTCCGTCGCAAAAGGGTCAAGCCCTCCAGCCCGCC 2171
Dh 2515 GTGGACAGGATGAAGAAACAGAGCTTCCGTCGCAAAAGGGTCAAGCCCTCCAGCCCGCC 2574
Qy 2172 CCGTCCCTGCACTCTGCTCTCTCCAGTTCATCTCTGCAAGCCAGCCAGGAGCAAC 2231
Dh 2575 CCGTCCCTGCACTCTGCTCTCTCCAGTTCATCTCTGCAAGCCAGCCAGGAGCAAC 2634
Qy 2232 CAGTGGCCCCAAGGAGGAGAGATCTCTAGGAGCCAGCCGCGGAGGCTGGAGGGCT 2291
Dh 2635 CAGTGGCCCCAAGGAGGAGAGATCTCTAGGAGCCAGCCGCGGAGGCTGGAGGGCT 2694
Qy 2292 GCGAGATGCTTCCCTCATCCACCTCCAGCGGTCCAGGTCTTGTGCTGTCCAGAGC 2351
Dh 2695 GCGAGATGCTTCCCTCATCCACCTCCAGCGGTCCAGGTCTTGTGCTGTCCAGAGC 2754
Qy 2352 TCCGTGACACACAGCCAGATCAGAGGAGCAGAGCCAGAGATAGTCTTTTGTGCT 2411
Dh 2755 TCCGTGACACACAGCCAGATCAGAGGAGCAGAGCCAGAGATAGTCTTTTGTGCT 2814
Qy 2412 TTCTGGCCTCTGGCTAGTCACTTTTTCATATGCGCTTACAGATGTGCTTTGCTAGAA 2471
Dh 2815 TTCTGGCCTCTGGCTAGTCACTTTTTCATATGCGCTTACAGATGTGCTTTGCTAGAA 2874
Qy 2472 ATAAACACATTTTCATATTTGTTA 2497
Dh 2875 ATAAACACATTTTCATATTTGTTA 2900

RESULT 4
BC004176
LOCUS BC004176 1905 bp mRNA linear PRI 12-Jul-2001
DEFINITION Homo sapiens, similar to hypothetical protein FLJ10928, clone
ACCESSION BC004176
VERSION BC004176.1 GI:13278815

KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1905)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-rt@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kelleman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 2 Row: m Column: 8.
location/Qualifiers
1. 1905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:2772 IMAGE:2958967"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
272..1171
/codon_start=1
/product="Similar to hypothetical protein FLJ10928"
/protein_id="AAH04176.1"
/db_xref="GI:33278816"
/translation="MAREIDNYPFRFFTHNVLWDESAQLPHEKTHRTIFARA
QGTIVLHCKMVGYSRATVLAAMQYECLEALHVDLRIARINPGLLOI
YQILITLRSRSHWQKGVSPVPEHAPSTFPLPEPEEGEGEKVGMESGA
APKEPGRPRINRGYMRSLISLEPSTLESTEDMPEVSHSSHEPPLQPP
QIARTKGGQVDRGPRPALKSRSQSVTLQGSAAVYANRTQAFQEGEQQGGGQGPCTIS
STPRKRVVHQASVHDSGESEA".
BASE COUNT 418 a 627 c 517 g 343 t
ORIGIN
Query Match 64.6%; Score 1641; DB 9; Length 1905;
Best Local Similarity 94.3%; Pred. NO. 0;
Matches 197; Conservative 0; Mismatches 5; Indels 103; Gaps 5;
QY 712 GAGCAGCGGATCCGCTGAGCTGTGAAAGTGTGATGTCAGTGGAGAGTGTG 771
DB 6 GAGAGGGGATCCGCTGCT-ACGTGTGAAAGTGTGATGTCAGTGGAGAGTGTG 64
QY 772 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC 831
DB 65 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC 124
QY 832 CGTGACTCATGCAACACATGCTGCTGCTGGTGCGACAGCGGGAGCGAGCTCCCGC 891
DB 125 CGTGACTCATGCAACACATGCTGCTGCTGGTGCGACAGCGGGAGCGAGCTCCCGC 184
QY 892 ATCTTCCCACTTCTACTGGGCTCAAGTGAAGCAGCAAACTGGAGAGCTGCAG 951
DB 185 ATCTTCCCACTTCTACTGGGCTCAAGTGAAGCAGCAAACTGGAGAGCTGCAG 244
QY 952 AGGAACAGGGTACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCAACCTGAG 1011
DB 245 AGGAACAGGGTACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCAACCTGAG 304

QY 1012 CGCTTACCTACACACATGTGCGCCTGTGGATGAGAGATGCGGCCACCTGTGCGGAC 1071
DB 305 CGCTTACCTACACACATGTGCGCCTGTGGATGAGAGATGCGGCCACCTGTGCGGAC 364
QY 1072 TGGAGAGAGACGACCGCTTCATTATGAGCTTCCAGAGACAGAGGACCCACGTGTGTC 1131
DB 365 TGGAGAGAGACGACCGCTTCATTATGAGCTTCCAGAGACAGAGGACCCACGTGTGTC 424
QY 1132 CACTTCAGATGGGGTCAAGCCGCTCAGCGGCCACAGTGTGCGCTATGCCATGAACGAG 1191
DB 425 CACTTCAGATGGGGTCAAGCCGCTCAGCGGCCACAGTGTGCGCTATGCCATGAACGAG 484
QY 1192 TACGAATCAGCTGAGCAGAGCCCTGCGCCACAGCTGAGAGACTCGGCCATCGCCGC 1251
DB 485 TACGAATCAGCTGAGCAGAGCCCTGCGCCACAGCTGAGAGACTCGGCCATCGCCGC 544
QY 1252 CCCAACCTTGCGCTTCCTGCGCCAGCTGAGATTTACAGAGGCAATCCTGAGC----- 1302
DB 545 CCCAACCTTGCGCTTCCTGCGCCAGCTGAGATTTACAGAGGCAATCCTGAGC----- 604
QY 1303 ----- 1302
DB 605 CAGAGCATGTCGTGGAGACAGAAAGTGGGTGGGTCTCCCAAGAGAGACCCACGCCCT 664
QY 1303 -----GCCAGAACCTGAGGGGTGGGGAGAGAG 1333
DB 665 GAAGTCTTACACACTTCCCACTTTCCTCCGACAGAACTGAGGGTGTGGGAGAGAGAG 724
QY 1334 GTTGTAGGATGAG 1393
DB 725 GTTGTAGGATGAG 782
QY 1394 GTATTAACCTCCGAGAGGCTCATGAGTCCATCAGTCTTGTGAGGCCCTTGGAGAGCTG 1453
DB 783 GTATTAACCTCCGAGAGGCTCATGAGTCCATCAGTCTTGTGAGGCCCTTGTGAGAGCTG 841
QY 1454 GAGAGACCTCGTAGGACAGGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
DB 842 GAGAGACCTCGAG 900
QY 1514 TGAAGAGCCTGTGAG 1573
DB 901 TGAAGAGCCTGTGAG 960
QY 1574 CAGGGGCTCAGCCTGAGCTGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1633
DB 961 CAGGGGCTCAGCCTGAGCTGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1634 CGTGTGGCCAAACCGAG 1693
DB 1021 CGTGTGGCCAAACCGAG 1080
QY 1694 GGGAGAGCCTGCAATTTCTCTACGCCAGGTTCCGGAAGTGTGAGAGAGAGAGAGAGAG 1753
DB 1081 GGGAGAGCCTGCAATTTCTCTACGCCAGGTTCCGGAAGTGTGAGAGAGAGAGAGAGAG 1140
QY 1754 GCATGACAGTGAAG 1813
DB 1141 GCATGACAGTGAAG 1200
QY 1814 ACTGAAGAGATCCCACTCTGTTGGAAGAAACACCTCAGCTGTGTCGGAGACATTC 1873
DB 1201 ACTGAAGAGATCCCACTCTGTTGGAAGAAACACCTCAGCTGTGTCGGAGACATTC 1260
QY 1874 CTCTAGCTCGGCCCATACCCGCTCACTACAGCTCAACCTCCACCTGTCTACAGGC 1933
DB 1261 CTCTAGCTCGGCCCATACCCGCTCACTACAGCTCAACCTCCACCTGTCTACAGGC 1320
QY 1934 CTCACTTCCACCTCTGTCTACAGCTCAACCTCTCAAGCTTTAAGTCCAGGCCAT 1993
DB 1321 CTCACTTCCACCTCTGTCTACAGCTCAACCTCTCAAGCTTTAAGTCCAGGCCAT 1380

QY 1994 GTCGCTGCTCAAGGCGCTCAAGACTTTCTAACGSGATGTGTAGAGGAGCTGAAGTA 2053
|||||
Db 1381 GTCGCTGCTCAAGGCGCTCAAGACTTTCTAACGSGATGTGTAGAGGAGCTGAAGTA 1440
QY 2054 CTTTGGGGGCAACAGCAGCCTAGTTTCACTTCACTAGGCTGACACTACCTGT 2113
|||||
Db 1441 CTTTGGGGGCAACAGCAGCCTAGTTTCACTTCACTAGGCTGACACTACCTGT 1500
QY 2114 GGCACGGAATGAAGAACAGAGCTTCCGTCGAAAGAGGATACCCCTCCAGCCGCC 2173
|||||
Db 1501 GGCACGGAATGAAGAACAGAGCTTCCGTCGAAAGAGGATACCCCTCCAGCCGCC 1560
QY 2174 CTCCTGCGACCTCTCTCTCTCTCCAGTTTCCTTCTGGAACAGCAGGCGAGCAACCA 2233
|||||
Db 1561 CTCCTGCGACCTCTCTCTCTCTCCAGTTTCCTTCTGGAACAGCAGGCGAGCAACCA 1620
QY 2234 GTGGCCCCCAAGAGGAGGAGGATCTCTAGGCCCCAGCCGGGAGGCTGGAAGGCGTGG 2293
|||||
Db 1621 GTGGCCCCCAAGAGGAGGAGGATCTCTAGGCCCCAGCCGGGAGGCTGGAAGGCGTGG 1680
QY 2294 CAGATGCTTCCCTCATCCACCTCCACCGGTCCAGGCTTTGGTCTGCTCCAGACCTC 2353
|||||
Db 1681 CAGATGCTTCCCTCATCCACCTCCACCGGTCCAGGCTTTGGTCTGCTCCAGACCTC 1740
QY 2354 CTGTGACACACGCGCAGATCAACAGGCGCAGGCGCAGAGATAGTCTTTTGTCTTT 2413
|||||
Db 1741 CTGTGACACACGCGCAGATCAACAGGCGCAGGCGCAGAGATAGTCTTTTGTCTTT 1800
QY 2414 CTGGCCTCTGCTAGTCACTTTTCTAGCCTTACGCTTGTGCTTGTGCTTGTACTGAAAT 2473
|||||
Db 1801 CTGGCCTCTGCTAGTCACTTTTCTAGCCTTGTGCTTGTGCTTGTGCTTGTACTGAAAT 1860
QY 2474 AAAACACATTTTCATTTTGTGTTAAAAA 2518
|||||
Db 1861 AAAACACATTTTCATTTTGTGTTAAAAA 1905

RESULT 5
BC004210 1905 bp mRNA linear PRI 12-JUL-2001
LOCUS
DEFINITION Homo sapiens, similar to hypothetical protein FLJ10928, clone
MGC:4436 IMAGE:2958967, mRNA, complete cds.
ACCESSION BC004210
VERSION BC004210.1 GI:13278902
KEYWORDS MGC.
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIR-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapubs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
<http://www.systemsbio.org>
contact: amadanes@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Keltman and Anuradha Madan

REMARK
COMMENT Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 10 Row: 9 Column: 1.
Location/Qualifiers
1. 1905

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:4436 IMAGE:2958967"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_id="NIR_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pORF7"
272. 1171
/codon_start=1
/product="Similar to hypothetical protein FLJ10928"
/protein_id="AAH04210.1"
/db_xref="GI:13278903"
/translation="MAREIDNVPPEFTYHNVRLWDEESADLLPHEKREHREIENARA
OETHLVHCKMGVSRSAATVLAAMKQYECSEALRIYVOELRIAPRPFELQLOI
YQGLILASROSHWEOKYGVSPHEHPAEPSTPPPLPPEEGEGEVVEMESOA
APKEBPDRPRINLEGVNRSISLEPSELESTSETPSMPEVSSHESEHPLOPFP
QLATPKGOOVDROGPALPKRSQSVYTIQGSANVANKRQARQEDBQGGGCEPCIS
SIPPRKRYVRQASVHDSGESEBA"
BASE COUNT 418 a 627 c 517 g 343 t
ORIGIN

Query Match 64.6%; Score 1641; DB 9; Length 1905;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 5; Indels 103; Gaps 5;

QY 712 GAGCAGGCGATCCGTCGCTGAGCTGTGGAAGGTTGATGTCAGTGAGAGTGTG 771
|||||
Db 6 GAGGAGGCGATCCGTCGCTGAGCTGTGGAAGGTTGATGTCAGTGAGAGTGTG 64
QY 772 ACTTCGAAGAGATCGCGCAGGCTCTGAGAGTGGCGCTGGGGGCTCCCTCCAGAGTAC 831
|||||
Db 65 ACTTCGAAGAGATCGCGCAGGCTCTGAGAGTGGCGCTGGGGGCTCCCTCCAGAGTAC 124
QY 832 CGTGACTTCATCGACACAGATGCTGCTGTGTGTGACACAGCGGAGCGAGCTCCCGC 891
|||||
Db 125 CGTGACTTCATCGACACAGATGCTGCTGTGTGTGACACAGCGGAGCGAGCTCCCGC 184
QY 892 ATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAAGAGGAGGAGAACTGTGAGAGTGTGAG 951
|||||
Db 185 ATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAAGAGGAGGAGAACTGTGAGAGTGTGAG 244
QY 952 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATTGAACATCTTACCTTGAG 1011
|||||
Db 245 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATTGAACATCTTACCTTGAG 304
QY 1012 CGCTTACCTTACCAATGTGCGCTCTGGGATGAGAGTGGCCCACTCTGCGCAC 1071
|||||
Db 305 CGCTTACCTTACCAATGTGCGCTCTGGGATGAGAGTGGCCCACTCTGCGCAC 364
QY 1072 TGAAGAGAGCGACCGCTTCAATTTGAGGCTCAAGAGACAGGCGCACCGACCTGTGGTC 1131
|||||
Db 365 TGAAGAGAGCGACCGCTTCAATTTGAGGCTCAAGAGACAGGCGCACCGACCTGTGGTC 424
QY 1132 CACGTGCAAGATGGGCGCTACGCGCTCAAGCGGCAACAGTGTGCTTATGCGCATGAACAG 1191
|||||
Db 425 CACGTGCAAGATGGGCGCTACGCGCTCAAGCGGCAACAGTGTGCTTATGCGCATGAACAG 484
QY 1192 TACGAATGCAAGCTGGAGAGGCGCTGGCGACGTGAGAGAGTCTCCGGCCATGTGCCCGC 1251
|||||
Db 485 TACGAATGCAAGCTGGAGAGGCGCTGGCGACGTGAGAGAGTCTCCGGCCATGTGCCCGC 544
QY 1252 CCGAACCCTGGCTTCTGCGGCAAGCTCAGATCTACAGAGGAGCTCTGAGC----- 1302
|||||
Db 545 CCGAACCCTGGCTTCTGCGGCAAGCTCAGATCTACAGAGGAGCTCTGAGC----- 604
QY 1303 ----- 1302
Db 605 CAGAGCCATGTCTGGAGAGAGAAAGTGGTGGGCTCCCGACAGAGAGCACCCAGCCCT 664
QY 1303 -----GCCAGACCTGAGGCTGTGGGAGGAGAG 1333
Db 665 GAGTCTCTACACCATTTCCACCTCTTCCGCGACAGACCTGAGGCTGTGGGAGAGAG 724

Db 541 TTCAGCGTGCAGCTGTGGGGCAAGCCGATCTTCAAGCCATCTCCATCCAGACCATTG 600
 Qy 688 T----- 688
 Db 601 TGGGCCACACTCCAGGTATTGACCAAGCATGTGAGGAGACTCTAGAGCAGCGGCTTGA 660
 Qy 689 ----- 688
 Db 661 CCGGGTGCAGTGGCCTCACCTGGGCAAGCCACTACAGAGAGACTGAACTCCAGACAG 720
 Qy 689 ----- 688
 Db 721 AACTGCTCAATGATGAGCGGCTATGGCCGACTGAGTCTCTGGGCTCCAGCGCC 780
 Qy 689 -----GGTCCCTCAGAACAGAGAGAGAGATGAGAGGGGATCCGCTGAGCTGG 738
 Db 781 GAGCTGGGCGGCTCCTCAGAACAGAGAGAGATGAGAGCGAGGATCCGCTGAGCTGG 840
 Qy 739 AAGGTGTGATGTAGTACGACCTGGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTTG 798
 Db 841 AAGGTGTGATGTAGTACGACCTGGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTTG 900
 Qy 799 GAGCTGCCTGGGGCTCCCTCCAGCATGACGTTGACTTCAATGACAGACATGCTG 858
 Db 901 GAGCTGCCTGGGGCTCCCTCCAGCATGACGTTGACTTCAATGACAGACATGCTG 960
 Qy 859 CTGCTGGTGCACAGCGGGACCGAGCTCCGATCTTCCCACTTCACTTACCTGGGCTCA 918
 Db 961 CTGCTGGTGCACAGCGGGACCGAGCTCCGATCTTCCCACTTCACTTACCTGGGCTCA 1020
 Qy 919 GAGTGAACGAGCAAACTGTGAGAGCTGAGAGGAAAGAGGCTCACCCACATCTTGAAC 978
 Db 1021 GAGTGAACGAGCAAACTGTGAGAGCTGAGAGGAAAGAGGCTCACCCACATCTTGAAC 1080
 Qy 979 ATGGCCCGGGAGATTGCAACTTCTACCTGTAGAGCTTCACTTCACTTCACTTGGCGCTC 1038
 Db 1081 ATGGCCCGGGAGATTGCAACTTCTACCTGTAGAGCTTCACTTCACTTCACTTGGCGCTC 1140
 Qy 1039 TGGGATGAGAGTTCGCGCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 Db 1141 TGGGATGAGAGTTCGCGCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 Qy 1099 GGTGCAAGAGACAGAGGACCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAGAG 1158
 Db 1201 GGTGCAAGAGACAGAGGACCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAG 1260
 Qy 1159 GCGGCCACAGTGTGCGCTATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
 Db 1261 GCGGCCACAGTGTGCGCTATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Qy 1219 GCGCACGTGAGAGAGCTCGGCGCATCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
 Db 1321 GCGCACGTGAGAGAGCTCGGCGCATCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 Qy 1279 CAGATCTACAGAGGATCTGACG----- 1302
 Db 1381 CAGATCTACAGAGGATCTGACGAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Qy 1303 ----- 1302
 Db 1441 GGTGGGCTCTCCCGAG 1500
 Qy 1303 --GCCAGAACTGAG 1360
 Db 1501 CCGCCAGAACTGAG 1560
 Qy 1361 GCGCCGAAAG 1420
 Db 1561 GCGCCGAAAG 1618
 Qy 1421 CCATGAGCTTCTGAG 1480
 Db 1619 CCATGAGCTTCTGAG 1676

Qy 1481 GCGAGAGCTTCTCTTCCACAGAGTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 1540
 Db 1677 GCGAGAGCTTCTCTTCCACAGAGTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 1736
 Qy 1541 GCTTGCAGGAG 1600
 Db 1737 GCTTGCAGGAG 1796
 Qy 1601 CCGCCAGTCAAGTGTATCCCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1660
 Db 1797 CCGCCAGTCAAGTGTATCCCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1856
 Qy 1661 CAG 1720
 Db 1857 CAG 1916
 Qy 1721 CAGGTTCCGAG 1780
 Db 1917 CAGGTTCCGAG 1976
 Qy 1781 CTGA 1784
 Db 1977 CTGA 1980
 RESULT 7
 AX398768 1755 bp DNA linear PAT 27-MAY-2002
 LOCUS AX398768
 DEFINITION Sequence 4 from Patent WO0220732.
 ACCESSION AX398768
 VERSION AX398768.1 GI:21261303
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Liou,J.R.
 TITLE Regulation of human map kinase phosphatase-like enzyme
 JOURNAL Patent: WO 0220732-A 4 14-MAR-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source location/Qualifiers
 1..1755.
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 350 a 377 c 505 g 323 t
 ORIGIN
 Query Match 51.3%; Score 1303.6; DB 6; Length 1755;
 Best Local Similarity 94.2%; Pred. No. 2,4e-258;
 Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;
 Qy 979 ATGGCCCGGGAGATTGCAACTTCTACCTGAGCGCTTACCTACCAATGTGGGCTTC 1038
 Db 322 ATGGCCCGGGAGATTGCAACTTCTACCTGAGCGCTTACCTACCAATGTGGGCTTC 381
 Qy 1039 TGGGATGAGAGAGTGGCGCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAG 1098
 Db 382 TGGGATGAGAGAGTGGCGCCAGCTGTGCGGACGAGAGAGAGAGAGAGAGAGAGAG 441
 Qy 1099 GCTGCAAGAGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
 Db 442 GCTGCAAGAGAGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
 Qy 1159 GCGGCACAGTGTGGGCTTATGCAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
 Db 502 GCGGCACAGTGTGGGCTTATGCAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561
 Qy 1219 GCGCACGTGCAAG 1278
 Db 562 GCGCACGTGCAAG 621

OY	1039	TEGGAGTAAAGGAAATCCGGCCCAAGTGTGTCGCCAGCTGGAAGGAGACGACACCGCTTCAATTAG	1098
Db	382	TEGGATGAGGAATCGGCCCAAGCTGTCTGCGGCACTGTGAAGGAGACGACACCGCTTCAATTAG	441
OY	1099	GCTCGAAGAGCAGAGGGACCCACGCTGCTGGGTCCTCAGATGAGGGGTAGGCGCTCA	1158
Db	442	GCTCGAAGAGCAGAGGGACCCACGCTGCTGGTCCACTCAGTCAAGATGAGGGGTAGGCGCTCA	501
OY	1159	GGGGCCACAGTGTGTGGCTTATGCCATGAAGCAAGTACGAATGCASCCTTGAGACAGGCTTG	1218
Db	502	GGGGCCACAGTGTGTGGCTTATGCCATGAAGCAAGTACGAATGCAGCTTGAGACAGGCTTG	561
OY	1219	CGCCACGTGCAGAGGACTCGGGCCCATGCGCCCGCCGCCCCAACCTGGCTTCTGTGCGCAAGCTG	1278
Db	562	CGCCACGTGCAGAGGACTCGGGCCCATGCGCCCGCCGCCCCAACCTGGCTTCTGTGCGCAAGCTG	621
OY	1279	CAGATCTACCAAGGGGATCTGTACGGGCCAAGAACTGAGAGGGTGTGGGGAGAGAAAGTTGT	1338
Db	622	CAGATCTACCAAGGGGATCTGTACGGGCCAAGAACTGAGAGGGTGTGGGGAGAGAAAGTTGT	681
OY	1339	AGGATGGAAAGAGGCCAGAGGACGCCGAAAGAAAGACCTGTGGGGCCAGGGGGCAAGCTATA	1398
Db	682	AGGATGGAAAGAGGCCAGAGGACGCCGAAAGAAAGACCTGTGGGGCCAGGGGGCAAGCTATA	739
OY	1399	AACCTCCGAGGGGTCATAGAGGTCCATCAGTCTTGTGAGGCTCTCTTGGAGAGGTGGAGAG	1458
Db	740	AACCTCCGAGGGGTCATAGAGGTCCATCAGTCTTGTGAGGCTCTCTTGGAGAGGTGGAGAG	798
OY	1459	CACCTCGATGACACAGTACATGCGACAGGTCTTCTCTCCACGAGCTTTCACATGAAG	1518
Db	799	CACCTCGAGACACAGTACATGCGACAGGTCTTCTCTCCACGAGCTTTCACATGAAG	857
OY	1519	AGCCTCGACGCCCTTCCACAGCTTGTCAAGAGCAACAAAGGAGGCGACAGGTGGACAGAGG	1578
Db	858	AGCCTCGACGCCCTTCCACAGCTTGTCAAGAGCAACAAAGGAGGCGACAGGTGGACAGAGG	917
OY	1579	GGGCTCAGGCTGCGCTGGAATCCCGGCATGCTCACTGTTACCTCCAGGGGCAAGTCCGCTGG	1638
Db	918	GGGCTCAGGCTGCGCTGGAATCCCGGCATGCTCACTGTTACCTCCAGGGGCAAGTCCGCTGG	977
OY	1639	TGGCCAAACCGGACCCAGGCTTTCAGAGAGCAGAGGAGAGGGGCGAGGGGCGAGGGAG	1698
Db	978	TGGCCAAACCGGACCCAGGCTTTCAGAGAGCAGAGGAGAGGGGCGAGGGGCGAGGGAG	1037
OY	1699	AGCCCTGCATTTCTCTTACGCCAGGTTCCGGAAGGTGTGTGAGACAGGCGACGTCGATG	1758
Db	1038	AGCCCTGCATTTCTCTTACGCCAGGTTCCGGAAGGTGTGTGAGACAGGCGACGTCGATG	1097
OY	1759	ACAGTGAAGAGGAGGGCGAGGCGCTGAGGCTTCACACATGCCACGCTCCCTGACACTGA	1818
Db	1098	ACAGTGAAGAGGAGGGCGAGGCGCTGAGGCTTCACACATGCCACGCTCCCTGACACTGA	1157
OY	1819	AGAGGATCCAAACTCTCTTGTGAAGAAACACCCCTCAGGTGTGTGGCGCAACACTTCTCTGC	1878
Db	1158	AGAGGATCCAAACTCTCTTGTGAAGAAACACCCCTCAGGTGTGTGTGGCGCAACACTTCTCTGC	1217
OY	1879	AGCTCGGCCCCATACCCGCTACATAGAGCTCACCTCCACCCGCTGTCACTAGGCGCTCAC	1938
Db	1218	AGCTCGGCCCCATACCCGCTACATAGAGCTCACCTCCACCCGCTGTCACTAGGCGCTCAC	1277
OY	1939	CTGCCACCCCTGTCACTACAGGCTCACCTCTACAGCTTAAGTCCAGGGGCGAATGTGTG	1998
Db	1278	CTGCCACCCCTGTCACTACAGGCTCACCTCTACAGCTTAAGTCCAGGGGCGAATGTGTG	1337
OY	1999	CTGTGCAAGGGGCTCAACACTTTTCTAACCTGGGATGTGTGTAGAGGACTGAAGGTACCTTT	2058
Db	1338	CTGTGCAAGGGGCTCAACACTTTTCTAACCTGGGATGTGTGTAGAGGACTGAAGGTACCTTT	1397
OY	2059	GGGGGCAACAGACCCCTGTGTTCAATTCTCAACTGTAGGCCCTGCACACTCACTGTGGCAC	2118
Db	1398	GGGGGCAACAGACCCCTGTGTTCAATTCTCAACTGTAG-----1434	1434

QY	2119	GGAAATGAAAACGAGACTCTCCGCTGCAAAAAGGCTCAGCGCTCCCAACCCCGCCCCCTCCC	2178
Db	1435	-----CCC	1437
QY	2179	TGCACCTCTGCTCCTCTCCAGTTCAATTCCTGSAACGACCGAGGCCAAGCACTCAGTGGC	2238
Db	1438	TGCACCTCTGCTCCTCTCCAGTTCAATTCCTGSAACGACCGAGGCCAAGCACTCAGTGGC	1497
QY	2239	CCCCAAAGCAGGCGAGATCCTCAGGCCCCAGCCGCGGGAGGCTGSAAGGAGCTGGCGAGAT	2298
Db	1498	CCCCAAAGCAGGCGAGATCCTCAGGCCCCAGCCGCGGGAGGCTGSAAGGAGCTGGCGAGAT	1557
QY	2299	CGCTTCCCTCATTCACCTCCACACGCGTGCAAGCTTCTGCTGCTGTGCCAGACCTCTGTTG	2358
Db	1558	CGCTTCCCTCATTCACCTCCACACGCGTGCAAGCTTCTGCTGCTGTGCCAGACCTCTGTTG	1617
QY	2359	ACACACAGCCCGAGATCACAGGGCACCAGGCCAGAGATAGTCTTCTTTTGTCTTCTTGCGC	2418
Db	1618	ACACACAGCCCGAGATCACAGGGCACCAGGCCAGAGATAGTCTTCTTTTGTCTTCTTGCGC	1677
QY	2419	CTGTGGCTAGCAGATTTTTCATATACCTTAAGATATCGGCTTGTGCTACTGAGAAATAAAG	2478
Db	1678	CTGTGGCTAGCAGATTTTTCATATACCTTAAGATATCGGCTTGTGCTACTGAGAAATAAAG	1737
QY	2479	ACATTTTCATATTTGGTT	2496
Db	1738	ACATTTTCATATTTGGTT	1755

RESULT 9				
LOCUS	AX398774			
DEFINITION	AX398774	2322 bp	DNA	linear
ACCESSION	Sequence 10 from Patent WO0220732.			PAT 27-MAY-2002
VERSION	AX398774			
KEYWORDS	AX398774.1	G1:21261307		
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE				
AUTHORS	Liou,J.R.			
TITLE	Regulation of human map kinase phosphatase-like enzyme			
JOURNAL	Patent: WO 0220732-A 10 14-MAR-2002;			
	Bayer Aktiengesellschaft (DE)			
FEATURES	Location/Qualifiers			
source	1..2322			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	467 a 751 c 718 g 386 t			
ORIGIN				

Query Match	Local Similarity	50.0%	Score 1270.6	DB 6	Length 2322
Best Local Similarity	82.0%	Pred. No. 1.6e-251			
Matches 1794	Conservative	0	Mismatches 79	Indels 316	Gaps 16
QY	126	CGGCGCTCCACGCCCGCTGGGGCCCTGGGAGCCAGGCGGTCCAGCCAGAGGATCGACTCCA	185		
Db	147	CGCCCTCACCCCTGGGGGCTGTCTCTCTCGGACGACCTGGTCCAGGAGGATCGACTCCA	206		
QY	186	GGGAAAGCGAGACTTTGGCGTGTCC - CTGGGGCGTGTCTCTGGGACTCGAGATGGAGGG	244		
Db	207	CGGA ---- AGAGCTTTGGGTCTCTCTGGGGCGTGTCTCTGGGACTCGAGATGGAGGG	262		
QY	245	ACATGATGATTCACCCAGAGGCCACTTCTGAGCCAAACAGAGAGGCCCGGATGAGCAGG	304		
Db	263	ACAAATGATGATTCACCAAGAGGCCACTTCTGAGCCAAACAGAGAGGCCCGGATGAGCAGG	322		
QY	305	AGCTTCACGCGG --- ACAGAGAGACTT - CGGGAGAGATATCCAGAGTCCCCAGAGCA	359		
Db	323	AGCTTCACGCGGAGAACCCAGACAGACTTTTCGATGAAAGATCCAGAGTCCCCAGAAACA	382		
QY	360	GGAGAGCAGAGGACGACACCTGCACCTTCATGATACAGCTGTGTGAGGCGGACGAGTGCAT	419		

Db 383 GGAGAGAGAGAGCAGCAATGCAACCTCATAGAGCGCTGTAGAGCCGACGAGGATGACAT 442
 Oy 420 CCGGCTGGCA--GGCCACTGGAGGACACCCGGCTTCCCGGCTCCGTACTGC----- 472
 Db 443 CCGACTTGGAGACCCAACTGGAGGACACCCGGGCTCCCGGGATCCGGATACCTTGCCT 502
 Oy 473 TGGAGTTTTCACAGAGAAAG-GAGAAAGTCTGAG-CCAGATGAGAGGATCTCTGGG 530
 Db 503 TGGAGTTTTCACAGAGAAAGGAGGAGTCTGAGCCAGATGAGAGGATCTCTGGG 562
 Oy 531 -CGTGAATTTCCCTGACAGCAGCTCCCGACCTGCACCCCTGGGCTGTCTTGGCCCTCT 589
 Db 563 ACGTGAATTTCCCTGACAGCAGCTCCCGACCTGCACCCCTGGGCTGTCTTGGCCCTCT 622
 Oy 590 GGAGTGAACCCAGAGTGTACTAGATGAGAGCGGGGCTTACAGCTGACGCTGTGGTGGC 649
 Db 623 GGAGTGAACCCAGAGTGTACTAGATGAGAGCGGGGCTTACAGCTGACGCTGTGGTGGC 682
 Oy 650 AAAGCCGAGTCTTCAAGCCCATCTCCATCCAGACCATGT----- 688
 Db 683 AAAGCCGAGTCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGATATTGC 742
 Oy 689 ----- 688
 Db 743 ACCAAGCATGTGAGGACAGCTTAGGACAGCGGCTTGTACGGGTGGAGTGCCCTCACCT 802
 Oy 689 ----- 688
 Db 803 GGGCCAGCACTACAGAGAGAGACTGAACCTCGAAGAGCTGCCCTCAATGAGTGAAGG 862
 Oy 689 -----GGTCTCAGAAC 700
 Db 863 CTATGGCCGAGCTGAGAGTCTTGGGGCTCCAGAGCCGACCTGGCGGTCTCTAGAAC 922
 Oy 701 AGGAGCAGATGAGAGCAGCGATCCGTGCTGAGCTGTGAAAGTGTGAGTCACTAGTACC 760
 Db 923 AGGAGCAATGAGAGCGGCGATCCGTGCTGAGCTGTGAAAGTGTGAGTCACTAGTACC 982
 Oy 761 TGGAGAGTGTCACTTCCAAAGAGATCCCGCAGGCTGTGAGTGTGGGCTGGGCTCCCC 820
 Db 983 TGGAGAGTGTCTTCCAAAGAGATCCCGCAGGCTGTGAGTGTGGGCTGGGCTCCCC 1042
 Oy 821 TCCAGCAGTACCTGTACTTCACTTCAGAACACAGATGCTGCTGTGTGACAGCGGAGCC 880
 Db 1043 TCCAGCAGTACCTGTACTTCACTTCAGAACACAGATGCTGCTGTGTGACAGCGGAGCC 1102
 Oy 881 GAGCCTCCGCACTTCCCAACCTCTACCTGGGCTGAGAGTGAACGCAACCAACTGG 940
 Db 1103 GAGCCTCCGCACTTCCCAACCTCTACCTGGGCTGAGAGTGAACGCAACCAACTGG 1162
 Oy 941 AGGAGCTCAGAGAAAGGAGTCAACCCACATCTTGAACATGCCCCGGGAGATTGCAACT 1000
 Db 1163 AGGAGCTCAGAGAAAGGAGTCAACCCACATCTTGAACATGCCCCGGGAGATTGCAACT 1219
 Oy 1001 TCTAACCTGAGAGCTTCACTTCACTTCAACCAATGTGGGCTGTGGATGAGAGTGGGCGCAGC 1060
 Db 1220 TCTAACCTGAGAGCTTCACTTCACTTCAACCAATGTGGGCTGTGGATGAGAGTGGGCGCAGC 1279
 Oy 1061 TGTGCGCGCACTGGAAGAGAGCAGCCGTTCACTTGAAGCTGCAAGAGCAGAGGACCC 1120
 Db 1280 TGTGCGCGCACTGGAAGAGAGCAGCCGTTCACTTGAAGCTGCAAGAGCAGAGGACCC 1339
 Oy 1121 ACGTGTGTGTCCACTGCAAGATGGGGCTCAGACCGCTCAGCGGCCACAGTGTCTGTG 1180
 Db 1340 ACGTGTGTGTCCACTGCAAGATGGGGCTCAGACCGCTCAGCGGCCACAGTGTCTGTG 1399
 Oy 1181 CCATGAAGCACTAGCAATGAGCTGAGAGCAGCGCTGCGGCACAGCTGAGAGAGCTCGGC 1240
 Db 1400 CCATGAAGCACTAGCAATGAGCTGAGAGCAGCGCTGCGGCACAGCTGAGAGAGCTCGGC 1459
 Oy 1241 CCATGCGCGCGCCCAACCTGTGGCTTCTGCGCAGCTGACAGTCTACAGGGAATCTCTGA 1300
 |||||||

Db 1460 CCATGCGCGCGCCCAACCTGTGGCTTCTGCGCAGCTGACAGTCTACAGGGAATCTCTGA 1519
 Oy 1301 CG----- 1302
 Db 1520 CGGCCAGCGCCAGAGCCATGTCTGGGAGCAGAAAGTGGGTGGGTCTCCGAGAGAGC 1579
 Oy 1303 -----GCCAAGACCTGAGGGTGGT 1322
 Db 1580 ACCGAGCCCTGAAGTCTCTACACCATTTCCACCTCTCCGCGCAAGACCTGAGGGTGGT 1639
 Oy 1333 GGGAGAGAGAGTGTAGCATGAGAGAGCCAGGAGCCGCCAAGAGAGAGCTGGGG 1382
 Db 1640 GGGAGAGAGAGTGTAGCATGAGAGAGCCAGGAGCCGCCAAGAGAGAGCTGGGG- 1698
 Oy 1383 CCAGGGGACAGTATTAACCTCCGAGGGGTGATGAGTGCATACATCTTGTGAGACCTTC 1442
 Db 1699 -CCAGGGGACAGTATTAACCTCCGAGGGGTGATGAGTGCATACATCTTGTGAGACCTTC 1757
 Oy 1443 CTTGAGAGCTGAGAGCAGCTCAGTAGACAGTGAATGACAGAGAGTCTTCTTCCGAC 1502
 Db 1758 CTT-GGAGCTGAGAGAGCAGCTCAG-AGACAGTGAATGACAGAGAGTCTTCTTCCGAC 1815
 Oy 1503 GAGTCTTCAATGAAGAGCCTCTGAGCCCTTCCACAGCTTGAAGAGCAGAGGAGGC 1562
 Db 1816 GAGTCTTCAATGAAGAGCCTCTGAGCCCTTCCACAGCTTGAAGAGCAGAGGAGGC 1875
 Oy 1563 CAGCAGGTGAGAGAGGGGCTCAGCTTCCCTGAAAGTCTCCGCACTAGTGTATACCTTC 1622
 Db 1876 CAGCAGGTGAGAGAGGGGCTCAGCTTCCCTGAAAGTCTCCGCACTAGTGTATACCTTC 1935
 Oy 1623 CAGGAGCTGCGGTGTGGGCAACCGGACCGAGGCTTCCAGAGAGCAGAGAGGGGAG 1682
 Db 1936 CAGGAGCTGCGGTGTGGGCAACCGGACCGAGGCTTCCAGAGAGCAGAGAGGGGAG 1995
 Oy 1683 GGGCAGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCAGGTTCCGAGAGGTGTGAGA 1742
 Db 1996 GGGCAGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCAGGTTCCGAGAGGTGTGAGA 2055
 Oy 1743 CAGGCGACGTGATGAGAGTGGAGAGGGCGAGGCTGACGCCCTACACATCTCCAC 1802
 Db 2056 CAGGCGACGTGATGAGAGTGGAGAGGGGCGAGGCTGACGCCCTACACATCTCCAC 2113
 Oy 1803 GCTCCCTGACACTGGAAGAGATCCACAACTCTTGGAGAAACACCTCTCAGCTGTGTTC 1862
 Db 2114 GCTCCCTGACACTGGAAGAGATCCACAACTCTTGGAGAAACACCTCTCAGCTGTGTTC 2173
 Oy 1863 CGCACATTTCTCTCAGCTCGGCCCATATCCGCTACTACAGCTCACTCCACCCCT 1922
 Db 2174 CGCACATTTCTCTCAGCTCGGCCCATATCCGCTACTACAGCTCACTCCACCCCT 2233
 Oy 1923 GTCACTAGGGGCTCACTCCGACCCCTGTCACTACAGCTCACTCCACCTTAACT 1982
 Db 2234 GTCACTAGGGGCTCACTCCGACCCCTGTCACTACAGCTCACTCCACCTTAACT 2293
 Oy 1983 CCAAGGCCCATGTGTGCTGTCCAAAGGCG 2011
 Db 2294 CCAAGGCCCATGTGTGCTGTCCAAAGGCG 2322

RESULT 10
 AK000522
 LOCUS AK000522 2064 bp mRNA linear PRI 22-FEB-2000
 DEFINITION Homo sapiens cDNA FLJ20515 fis, clone KAT09889.
 ACCESSION AK000522
 VERSION AK000522.1 GI:7020673
 KEYWORDS o1igo capping; fis (full insert sequence).
 SOURCE Homo sapiens signed-ringing cell carcinoma cell_line:KATO III cDNA to mRNA, clone, lib:KAT clone:KAT09889.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Matanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,


```
Dh 992 ---GTTGGAGCTGGAGACACCTCTAG-AGACAGTGAACATGCCAGAGCTCTTCTCTCC 1047
Oy 1501 ACAGATCTTACATGAAGAGGCTCTGACGCCCTTCCACAGCTTCCAGAGGACCAAGGGAG 1560
Dh 1048 ACAGATCTTACATGAAGAGGCTCTGACGCCCTTCCACAGCTTCCAGAGGACCAAGGGAG 1107
Oy 1561 GCCACAGAGTGGACAGAGGGGGCTCAGCCTGCCCTGAAAGTCCCGCCAGTCACTGCTTACC 1620
Dh 1108 GCCACAGAGTGGACAGAGGGGGCTCAGCCTGCCCTGAAAGTCCCGCCAGTCACTGCTTACC 1167
Oy 1621 TCCAGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Dh 1168 TCCAGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
Oy 1681 AGGAGGACAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Dh 1228 AGGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1287
Oy 1741 GACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
Dh 1288 GACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1347
Oy 1801 ACAGTCTCTTACACTGAAGAGGATCCACACTCTCTTGAAGAAACCTCTACGCTCTT 1860
Dh 1348 ACAGTCTCTTACACTGAAGAGGATCCACACTCTCTTGAAGAAACCTCTACGCTCTT 1407
Oy 1861 GCGGACACATCTCTCTGAGTCCGCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Dh 1408 GCGGACACATCTCTCTGAGTCCGCCCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1467
Oy 1921 CTGCTACTAGAGGCTCAGCTCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Dh 1468 CTGCTACTAGAGGCTCAGCTCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527
Oy 1981 GTCCAGAGGCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Dh 1528 GTCCAGAGGCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1587
Oy 2041 GGGAGTGAAGTACCTTTGGGGGGCAACAGCACCCCTGATTTCTCACTGCTGCTGCTGCTG 2100
Dh 1588 GGGAGTGAAGTACCTTTGGGGGGCAACAGCACCCCTGATTTCTCACTGCTGCTGCTGCTG 1647
Oy 2101 CACACTCAGCTGTGGCAGGGAATGAAACAGAGCTTCCGTCGCAAAAAGGCTCAGCCCTC 2160
Dh 1648 CACACTCAGCTGTGGCAGGGAATGAAACAGAGCTTCCGTCGCAAAAAGGCTCAGCCCTC 1707
Oy 2161 CCAGCCCGCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Dh 1708 CCAGCCCGCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1767
Oy 2221 GGGCAGGACAGTGGGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Dh 1768 GGGCAGGACAGTGGGCGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1827
Oy 2281 CTGGAAGGGCTGGCAGATGCTTCCCTCATCCACCTCCACCGGCTCAGGCTCTTCTGCTGCT 2340
Dh 1828 CTGGAAGGGCTGGCAGATGCTTCCCTCATCCACCTCCACCGGCTCAGGCTCTTCTGCTGCT 1887
Oy 2341 GTCCCAACACTCTCTGTACACACAGCCAGATCACAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
Dh 1888 GTCCCAACACTCTCTGTACACACAGCCAGATCACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1947
Oy 2401 CTTTGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Dh 1948 CTTTGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2007
Oy 2461 TGTACTGAGAATAAACAACATTTTTCATATTTGTTTAAAAAAGGAGGAGGAGGAGGAGGAGG 2517
Dh 2008 TGTACTGAGAATAAACAACATTTTTCATATTTGTTTAAAAAAGGAGGAGGAGGAGGAGGAGG 2064
```

RESULT 11
AX405828

```
LOCUS AX405828 2061 bp DNA linear PAT.14-JUN-2002
DEFINITION Sequence 243 from Patent WO0222660.
ACCESSION AX405828
VERSION AX405828.1 GI:21439095
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tang,Y.T., Liu,C., Zhou,P., Asundi,V., Zhang,J., Zhao,Q.A., Ren,F.,
Xue,A.J., Yang,Y., Wehrman,T. and Drmanac,R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 243 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..2061
/organism="Homo sapiens"
/db_xref="taxon:9606"
147..1331
CDS
/codon_start=1
/note="unnamed protein product"
/protein_id="CAD34935.1"
/db_xref="GI:21439096"
/translation="MALTVRSPPSGSGASTPVGPMQAVQRRSLQHRQFAVLRLG
VLGDDGDNDDAEASEPEKAPSEELGDDTDFQSGSPQKQOEORHILMY
OLRRPDDIRLAQLAEPRLRLYLIVSRBEBSGSDTTLVGLVDFPSSSPCT
LGLVPLMSPTQVYVLDGSGSVSGSGSRIFKPSIDTMMATQVLYQAEEALGSG
LVPGSALVTEASHIÖERLNSÖSCLENTEAMADLESLEPSPAEPEGSSÖDÖMÖALR
AELMKVLESTSESTDPEVFSHESHEBLOPFLQARTKGQÖVDRQPOPALKS
ROSVYTLQGSVAVNRTQAFQEOEGQÖGQÖGQISSTPFRVRVROASVHDSGEEG
EA"
BASE COUNT 415 a 672 c 605 g 369 t
ORIGIN
Query Match 47.9%; Score 1216.4; DB 6; Length 2061;
Best Local Similarity 74.6%; Pred. No. 2.3e-240;
Matches 1875; Conservative 0; Mismatches 126; Indels 513; Gaps 7;
Oy 1 CCTGCTCTCGCGGGTCCAGAGACTGCCCGGGGTTGAGAGGAAGGGCCGCTCCCGGCTGC 60
Dh 61 CCTGCTCTCGCGGGTCCAGAGACTGC-CGCGGGGTTGAGAGGAAGGGCCGCTCCCGGCTGC 119
Oy 61 CAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGTCAGAGTACGCTTCCGCCCCG 120
Dh 120 CAGCCAGGTGCTCGCGGCTGCTCATGAGCCCTGTCAGAGTACGAGCCGCTCCCGG 179
Oy 121 GGCAGCGGCGCTTCCACGCGCTGGGGCCCTGGGACAGGCGGTCCAGGAGAGTCTGA 180
Dh 180 GGCAGCGGCGCTTCCACGCGCGTGGGGCCCTGGGACAGGCGGTCCAGGAGAGTCTGA 239
Oy 181 CTCACGCAAGGACAGAGCTTTGCGGTGCTCCGTGGGGCTGTCTTGGAGTGCAGATGGA 240
Dh 240 CTCACGCAAGGACAGAGCTTTGCGGTGCTCCGTGGGGCTGTCTTGGAGTGCAGATGGA 299
Oy 241 GGGCAATGATGATGACAGCAGGCGCAGTTCTGAGCCCAAGAGAGGCCCCGAGTGA 300
Dh 300 GGGCAATGATGATGACAGCAGGCGCAGTTCTGAGCCCAAGAGAGGCCCCGAGTGA 359
Oy 301 GAGGAGCTCCAGGGGAGCCAGACAGACTTGGGCAAGAGATCCAGAGTCCCGCAGAGCAG 360
Dh 360 GAGGAGCTCCAGGGGAGCCAGACAGACTTGGGCAAGAGATCCCGCAGAGCAG 419
Oy 361 GAGGAGCAGAGGACAGCAGCTGATGATGACAGCTGCTAGAGCCCGCAGATGATC 420
Dh 420 GAGGAGCAGAGGACAGCAGCTGATGATGACAGCTGCTAGAGCCCGCAGATGATC 479
Oy 421 GCGCTGGCAGCCAGCTGGAGGACCCCGGCTCCCGGCTCGGCTACTGCTGGAGTT 480
Dh 480 GCGCTGGCAGCCAGCTGGAGGACCCCGGCTCCCGGCTCGGCTACTGCTGGAGTT 539
Oy 481 TCTACAGAGAGAGAGAGTGTGAGCCAGAGATGAGAGCGTCTCTGGGCGTGAATTC 540
```

Db 540 TCTACAGAGAGAGAGAGTCTGTGAGCCAGATGAGACGGTCTCTGAGGCTGATTTTC 599
QY 541 CTTGACAGAGCTCCCGAGCTGCACCCCTGGGCTTGGTCTTCCCTCTGAGTGTACCC 600
Db 600 CTTGACAGAGCTCCCGAGCTGCACCCCTGGGCTTGGTCTTCCCTCTGAGTGTACCC 659
QY 601 CAGGTACTAGATGAGAGAGGGGGCTTTCAGCGTGTGAGGCAAGGCGGATC 660
Db 660 CAGGTACTAGATGAGAGAGGGGGCTTTCAGCGTGTGAGGCAAGGCGGATC 719
QY 661 TTCAAGCCATCTCATTCACACCATGTGTCTTACAGACAGAGCAGATGAGAGCGG 720
Db 720 TTCAAGCCATCTCATTCACACCATGTGTCTTACAGACAGAGCAGATGAGAGCGG 762
QY 721 ATCCGTGTGAGTGTGAGAGTGTGGATGTCACTGACCTGGAGAGTGTCTTCCAA 780
Db 763 ----- 762
QY 781 GAGATCCGACAGGCTCTGAGAGCTGCGCTGGGCTCCCTCCAGAGTACCTGACTTC 840
Db 763 -----TATTGC 768
QY 841 ATGCACACAGATGTGTCTGTGTGAGACAGGGAGCGGAGCGCTCCGCTATTTCCC 900
Db 769 ACAGAGCATGTGAGGAGAGCTGTAGGACAGCGGCTTGTACCGGG----- 811
QY 901 CACCTTACTGAGGCTGAGAGTGAACGACCAACCTGAGAGAGTGTGAGAGGACAGG 960
Db 812 ----- 811
QY 961 GTCAACCCACATCTTGACATGAGCCCGGAGATTGACAACTTCAACCTGAGCGCTTACC 1020
Db 812 -----TGGACATGCGCTTACCTGTGGCGCCAGCCACGACAGAGAGACATGAAAC 857
QY 1021 TACCACATGTGGGCTCTGTGATGAGAGTGGCCAGCTGTGCGCCAGCTGAAAGAG 1080
Db 858 TCCGAACAGAGTGGCTCAATGA----- 880
QY 1081 ACCGACCGCTTCAATGAGGCTGCAAGAGCAGAGGCAACCCAGCTGTGTCTGACAGCAAG 1140
Db 881 ----- 880
QY 1141 ATGGGGGTACGCCGTACAGCGGCACAGTGTGGCTATGCCATGAACGATGCAATGC 1200
Db 881 ----- 880
QY 1201 AGCTTGAGAGCGCTCGCGCCACGTGAGAGCTCGCGCCATCGCCGCCCAACCT 1260
Db 881 -----GTGACGCGCTATGGCGCCAGCTGAGATCTCTGCGGCTCCAGCGCGCTGCG 935
QY 1261 GCGTTCCTGGGAGCTGAGATCTACAGAGGATCTCTGAGCGCCAGAACTGAGGGTGG 1320
Db 936 GGGTCTCTCAGAACAGAGAGATGAGAGCGGATTC-----GTGC 976
QY 1321 TGGGAGAGAGAGTGTGTAGGATGAGAGAGAGAGAGCGAGCCCGAAGAGAGCTGG 1380
Db 977 TGAAGTGTGGAAGT----- 991
QY 1381 GGGCAGGGGACAGTAACTCCGAGGGGTGATGAGTGCATGATCTTCTGAGCC 1440
Db 992 ----- 991
QY 1441 TCCCTGGAGCTGAGAGACCTCAGTAGACAGTGAATGCAAGAGTCTTCTTCCC 1500
Db 992 -----GTTGAGAGCTGAGAGACCTCAG-AGACAGTGAATGCAAGAGTCTTCTTCCC 1047
QY 1501 ACAGAGTCTTACATGAAGAGCTTGTGAGAGCTTCCACAGATTGCAAGAGACAGGGAG 1560
Db 1048 ACAGAGTCTTACATGAAGAGCTTGTGAGAGCTTCCACAGATTGCAAGAGACAGGGAG 1107
QY 1561 GCCAGAGGTGAGACAGGGGGCTGAGCTGCCCTGAGTATCCCGCAGTCAAGTGTACCC 1620
Db 1108 GCCAGAGGTGAGACAGGGGGCTGAGCTGCCCTGAGTATCCCGCAGTCAAGTGTACCC 1167

QY 1621 TCCAGGCGATGCCGTGTGTGAGCCAAACCGAGCCCTTCCAGAGACAGAGAGGCG 1680
Db 1168 TCCAGGCGATGCCGTGTGTGAGCCAAACCGAGCCCTTCCAGAGACAGAGAGGCG 1227
QY 1681 AGGGGCGAGGCGAGGAGAGCCTTGCATTTCTCTACGCCAGGTTCCGGAAGTGTGA 1740
Db 1228 AGGGGCGAGGCGAGGAGAGCCTTGCATTTCTCTACGCCAGGTTCCGGAAGTGTGA 1287
QY 1741 GACAGCGCACGCTGATGAGAGAGAGAGAGGCGGCTGAGGCTTCCACATAGCCC 1800
Db 1288 GACAGCGCACGCTGATGAGAGAGAGAGAGGCGGCTGAGGCTTCCACATAGCCC 1347
QY 1801 ACCTCCCTGTGACATGAGAGAGATCCAACTCTTGGAGAAACACCTTCAAGTGTGT 1860
Db 1348 ACCTCCCTGTGACATGAGAGAGATCCAACTCTTGGAGAAACACCTTCAAGTGTGT 1407
QY 1861 GCGGACACATTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1408 GCGGACACATTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1467
QY 1921 CTGTCACTACGAGGCTTCACTCCACCCCTGTGACATGAGAGAGAGAGAGAGAG 1980
Db 1468 CTGTCACTACGAGGCTTCACTCCACCCCTGTGACATGAGAGAGAGAGAGAGAG 1527
QY 1981 GTCCAGGCGCATGTGTGCTGTGCAAGGAGTCAAGTCTTCACTGAGAGTGTGA 2040
Db 1528 GTCCAGGCGCATGTGTGCTGTGCAAGGAGTCAAGTCTTCACTGAGAGTGTGA 1587
QY 2041 GGGACTGAAGGTACCTTGTGGGGGCAACAGACCTTCACTGAGAGAGAGAGAGAG 2100
Db 1588 GGGACTGAAGGTACCTTGTGGGGGCAACAGACCTTCACTGAGAGAGAGAGAGAG 1647
QY 2101 CACACTACCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 1648 CACACTACCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1707
QY 2161 CCACCCCGCCCGCTCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 1708 CCACCCCGCCCGCTCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1767
QY 2221 GGCACAGGCAACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 1768 GGCACAGGCAACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1827
QY 2281 CTGGAAGGCGTGGAGATGCTTCCCTATCCACTCCAGGCTTCAAGGCTTGTGCT 2340
Db 1828 CTGGAAGGCGTGGAGATGCTTCCCTATCCACTCCAGGCTTCAAGGCTTGTGCT 1887
QY 2341 GTCCACAGAGCTCTGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Db 1888 GTCCACAGAGCTCTGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1947
QY 2401 CTTTGTGCTTCTTGTGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
Db 1948 CTTTGTGCTTCTTGTGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2007
QY 2461 TGTACTGAGAAATTAACATTTTCAATTTGTTTAAAAA 2514
Db 2008 TGTACTGAGAAATTAACATTTTCAATTTGTTTAAAAA 2061

RESULT 12
AX451362 1416 bp DNA linear PAT 03-JUL-2002
LOCUS
DEFINITION Sequence 20 from Patent W00224740.
ACCESSION AX451362
VERSION AX451362.1 GI:21698396
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

[illegible]

D	b	841	AAAGTGTGGATGCTAGTGACCTGGAGAGATGTCACCTTCCAAAGAGATCGCCAGGCTTG	900
Q	y	799	GAGCTGCAGCCTTG6GGCTCCCCCTCCAGACAGTACCCTCATCGACACACAGATGCTG	858
D	b	901	GAGCTGCAGCCTTG6GGCTCCCCCTCCAGACAGTACCCTCATCGACACACAGATGCTG	960
Q	y	859	CTGCTGTGTGACACACGGGAGCCGAGCTCCGCCATCTTCCCACACTCTTACCTGGGCTCA	918
D	b	961	CTGCTGTGTGACACACGGGAGCCGAGCTCCGCCATCTTCCCACACTCTTACCTGGGCTCA	1020
Q	y	919	GAGTGAACCGCACCAACCTGGAGAGCTGCAGAGAACAGGGTCAACCACATCTTTGAC	978
D	b	1021	GAGTGAACCGCACCAACCTGGAGAGCTGCAGAGAACAGGGTCAACCACATCTTTGAC	1080
Q	y	979	ATGGCCCGGGAGATTGACAACTTCTTACCCCTGAGCCCTTACCCTACACAATGTGGCCTC	1038
D	b	1081	ATGGCCCGGGAGATTGACAACTTCTTACCCCTGAGCCCTTACCCTACACAATGTGGCCTC	1140
Q	y	1039	TGGGATGAGGAGTCGCGCCACAGCTGCTGCGCACCTGGAAGAAGACGACCGCTCATTTAG	1098
D	b	1141	TGGGATGAGGAGTCGCGCCACAGCTGCTGCGCACCTGGAAGAAGACGACCGCTCATTTAG	1200
Q	y	1099	GCTGCAAGAGCACAGGGCACCCAGCTGCTGTGTTCCACTGCAAGATGGGCTAAGCGCTCA	1158
D	b	1201	GCTGCAAGAGCACAGGGCACCCAGCTGCTGTGTTCCACTGCAAGATGGGCTAAGCGCTCA	1260
Q	y	1159	GGGGCCACAGTCTGTGGCCTATGCAATGAAAGCAGTAGTACGATGAGCTGAGAGAGGCTTG	1218
D	b	1261	GGGGCCACAGTCTGTGGCCTATGCAATGAAAGCAGTAGTACGATGAGCTGAGAGAGGCTTG	1320
Q	y	1219	CGCCACGTGACAGGAGCTCCGGCCATCGCCCGCCCAACCTGGCTTCCTCGCCAGCTG	1278
D	b	1321	CGCCACGTGACAGGAGCTCCGGCCATCGCCCGCCCAACCTGGCTTCCTCGCCAGCTG	1380
Q	y	1279	CAGATCTACCAAGGAGTCTGTACGGCCAGAACCTTA	1314
D	b	1381	CAGATCTACCAAGGAGTCTGTACGGCCAGAACCTTA	1416
RESULT 13				
A	P002776			
L	OLOCUS	Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT		
D	EFINITION	SEQUENCE, 36 unordered pieces.		
A	AP002776			
A	AP002776.2	GI:12246853		
K	KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
S	SOURCE	Homo sapiens DNA, clone:RP11-126P21.		
O	ORGANISM	Homo sapiens		
R	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
A	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
T	TITLE	1 (bases 1 to 160903)		
J	JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,		
A	AUTHORS	Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
R	REFERENCE	Homo sapiens 160,903 genomic DNA of 11q		
A	AUTHORS	Published Only in Database (2000)		
T	TITLE	2 (bases 1 to 160903)		
J	JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,		
A	AUTHORS	Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
C	COMMENT	Direct Submission		
		Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical		
		and Chemical Research (RIKEN), Genomic Sciences Center (GSC);		
		1-7-22 Shutoho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan		
		(E-mail:hattori@gsc.riken.go.jp, ytk@choma.kanagawa.ac.jp),		
		Tel:81-45-503-9111, Fax:81-45-503-9170)		
		On Jan 16, 2001 this sequence version replaced gi:9188614.		
		----- Genome Center		
		Center: RIKEN Genomic Sciences Center(GSC)		
		Center code: RIKEN		
		Web site: http://hnp.gsc.riken.go.jp/		
		Contact: hattori@gsc.riken.go.jp		
		Project Information		

Db	7139	AGGTCCTCTTCCACGAGCTTTCACATGAAAGAGCCTCTGAGCCCTTCCACAGCTTG	7258
Qy	1546	CAAGGACCAAGGGAGCCAGCAGGTGAGCAGGGGGCTTCAGCCTGCGCTGAAGTCCGCC	1605
Db	7259	CAAGGACCAAGGGAGCCAGCAGGTGAGCAGGGGGCTTCAGCCTGCGCTGAAGTCCGCC	7318
Qy	1606	AGTCAGTGGTTACCTCCAGGGCAATGCCGTGTGTGGCCAAACCGGACCCAGGCTTTCAGG	1665
Db	7319	AGTCAGTGGTTACCTCCAGGGCAATGCCGTGTGTGGCCAAACCGGACCCAGGCTTTCAGG	7378
Qy	1666	AGCAGAGCAGGGGGCAGGGGGCAGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCAGGT	1725
Db	7379	AGCAGAGCAGGGGGCAGGGGGCAGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCAGGT	7438
Qy	1726	TCCGGAAAGTGGTGAGAGGCCACGCTGCATGACAGTGGAGAGGAGGGCGAGGCTGAG	1785
Db	7439	TCCGGAAAGTGGTGAGAGGCCACGCTGCATGACAGTGGAGAGGAGGGCGAGGCTGAG	7498
Qy	1786	CCCTCACACATGCGCCAGCCTCCCTCGACACATGAAGATCCACAATCTTTGGAGAAAC	1845
Db	7499	CCCTCACACATGCGCCAGCCTCCCTCGACACATGAAGATCCACAATCTTTGGAGAAAC	7558
Qy	1846	ACCCTCAGCTGTGTGGCCGACACATTCCTCTCAGACTCGGCCCCATACCCGTCACATACAG	1905
Db	7559	AACCCTCAGCTGTGTGGCCGACACATTCCTCTCAGACTCGGCCCCATACCCGTCACATACAG	7618
Qy	1906	CCTACACTCCACACCCCTGTACACTACGAGCCTACCTCCACACCCCTGTACACTACAGCCTCAC	1965
Db	7619	CCTACACTCCACACCCCTGTACACTACGAGCCTACCTCCACACCCCTGTACACTACAGCCTCAC	7678
Qy	1966	CTCCACACGCTTAAGTCCACAGGCCCATGTGCTGCCTGCACAAAGGGCTCAAGACTTTCTAA	2025
Db	7679	CTCCACACGCTTAAGTCCACAGGCCCATGTGCTGCCTGCACAAAGGGCTCAAGACTTTCTAA	7738
Qy	2026	CTGGGATGTGGTAGAGGAGCTGAAGTACCTTTGGGGCAACAGACCCCTAGTTTCATTC	2085
Db	7739	CTGGGATGTGGTAGAGGAGCTGAAGTACCTTTGGGGCAACAGACCCCTAGTTTCATTC	7798
Qy	2086	TCAACTCTAGCCCTGCACACTACCTGTGTGGCAGCGAATGAAACAGAGCTTCCGTGCAA	2145
Db	7799	TCAACTCTAGCCCTGCACACTACCTGTGTGGCAGCGAATGAAACAGAGCTTCCGTGCAA	7858
Qy	2146	AAAGGTCACGGCTCCACACCCCGCCCTCCCTGACACTCTCTGTCCTCCACAGTCAAT	2205
Db	7859	AAAGGTCACGGCTCCACACCCCGCCCTCCCTGACACTCTCTGTCCTCCACAGTCAAT	7918
Qy	2206	TCCTTGAACCAAGCCAGGCGCAAGCAACAGTGGCCCCCAAAGGCAGGAGATCTCAGGC	2265
Db	7919	TCCTTGAACCAAGCCAGGCGCAAGCAACAGTGGCCCCCAAAGGCAGGAGATCTCAGGC	7978
Qy	2266	CCCAGCCCCGGAGGCTGGAAAGGCTGTGGCAGATGCGCTTCCCTCATTCACCTCCACCGGTC	2325
Db	7979	CCCAGCCCCGGAGGCTGGAAAGGCTGTGGCAGATGCGCTTCCCTCATTCACCTCCACCGGTC	8038
Qy	2326	CAGGCTTTGCGTGCCTGCCAGACCTCTGTGACACACGCGCAATACACAGGGGACAG	2385
Db	8039	CAGGCTTTGCGTGCCTGCCAGACCTCTGTGACACACGCGCAATACACAGGGGACAG	8098
Qy	2386	GCCAGAGATAGTCTTTCTTTTGTCCCTTCTGTGGCCTGTGGCTAGTCAAGTTTTTCATAGCCT	2445
Db	8099	GCCAGAGATAGTCTTTCTTTTGTCCCTTCTGTGGCCTGTGGCTAGTCAAGTTTTTCATAGCCT	8158
Qy	2446	TACAGTATCTGGCTTTGTACTGAGAAATATAAAAACATTTTCAATTTGGTTA 2497	
Db	8159	TACAGTATCTGGCTTTGTACTGAGAAATATAAAAACATTTTCAATTTGGTTA 8210	

RESULT 15	AB072360	1416 bp	mRNA	linear	PRI 29-JAN-2002
LOCUS	AB072360				
DEFINITION	Homo sapiens mRNA for hssh-3, complete cds.				
ACCESSION	AB072360				
VERSION	AB072360.1	GI:18376668			

KEYWORDS	source	FEATURES
ORGANISM	Homo sapiens cDNA to mRNA.	
REFERENCE	Homo sapiens	
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
JOURNAL	Niwa, R., Nagata-Ohashi, K., Takeichi, M., Mizuno, K. and Uemura, T.	
MEDLINE	Control of actin reorganization by Slingshot, a family of phosphatases that dephosphorylate ADP/cofilin	
AUTHORS	Cell 108 (2), 233-246 (2002)	
JOURNAL	2 (bases 1 to 1416)	
REFERENCE	Niwa, R., Nagata-Ohashi, K., Hay, B. A., Takeichi, M., Mizuno, K. and Uemura, T.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular Genetics, The Institute for Virus Research, Kyoto University; Shogoin-Kawara-cho 53, Sakyo-ku, Kyoto 606-8507, Japan (E-mail: tuemura@virus.kyoto-u.ac.jp, Tel: 81-75-751-4031, Fax: 81-75-751-3989)	
FEATURES	Location/Qualifiers	
source	1..1416	
gene	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	1..1416	
CDS	/gene="hSSH-3"	
	1..1416	
	/gene="hSSH-3"	
	/standard_name="slingshot-3"	
	/note="a putative dual specific phosphatase"	
	/codon_start=1	
	/product="hSSH-3"	
	/protein_id="BAB84119.1"	
	/db_xref="gi:18376669"	
	/translation="MALVTYRSPSPGASHPVGPMDYAVORSRLOROSFAVLRALVLRGSDGNDNDVAESSEPLEKPSBELHGDITDQDSOSFQKEBQRIHLHWQLPRQODIRLAQLEAPRPRRLVLTSTREEGLSDQETVLGVDFPSSSCTGLALPLMSDQVLYLDDGGFSTVSGGQSRIFKPLISLTMMATLQVLAQCEALIGSGLVPGSALTMASHYQERLNSEOSLNTMTAMLDLESRLPPSAEPGSSSEDEOMQAIRALMKLVLDSDLESVTSKETRIQLMSKAWGSSSTVRSSTTRCCMMHSDPRASIRPPHLTSGSEMNANLEELORNRTHLLNMARELDNYPRLFTYHNVRLLDESSAQLPWHKETHRETEARAQGTTHLVHCKMVSNSAATVLAAMKQTECSLEQALRHVQELRPIA RPNQGFRLQIDYIGILTART"	
BASE COUNT	288 a 449 c 449 g 230 t	
ORIGIN		
Query Match	39.5%; Score 1003.4; DB 9; Length 1416;	
Best Local Similarity	86.5%; Pred. No. 1.9e-196;	
Matches 1225; Conservative	0; Mismatches 1; Indels 191; Gaps 3;	
QY	88 ATGGCCCTGTGTACAGTGTGCGCTTGCGCCCGGCGACGGGGCGCTCCACGCGCGTG66G 147	
DB	1 ATGGCCCTGTGTACAGTGTGCGCTTGCGCCCGGCGACGGGGCGCTCCACGCGCGTG66G 60	
QY	148 CCCTGGGACCAAGCGCGTCCAGCGCAAGGAGTGCATCCAGCGAAGGACAGAGCTTTGCGGTG 207	
DB	61 CCCTGGGACCAAGCGCGTCCAGCGAAGGAGTGCATCCAGCGAAGGACAGAGCTTTGCGGTG 120	
QY	208 CTCCTGGGGGCTGTCTCTGGGACTGTGAGATGGAGGGGACATGATGATGACAGAGGCC 267	
DB	121 CTCCTGGGGGCTGTCTCTGGGACTGTGAGATGGAGGGGACATGATGATGATGAGAGGCC 180	
QY	268 AGTTGTGAGCCCAAGCAAGAGGCCCGCGAGTAGAGGAGGAGTCCACGGGGACCAAGCAAC 327	
DB	161 AGTTGTGAGCCCAAGCAAGAGGCCCGCGAGTAGAGGAGGAGTCCACGGGGACCAAGCAAC 240	
QY	328 TTTCGGCGAAGATCCACAGTCCACAGAGAGCAGAGAGAGCAGAGGACACCTGCACCTC 387	
DB	241 TTTCGGCGAAGATCCACAGTCCACAGAGAGCAGAGAGAGCAGAGGACACCTGCACCTC 300	
QY	388 ATGTACAGCTGTCTGTAGAGCCGACGAGATGATCCGCTTGGAGCCCACTGGAGGACCC 447	
DB	301 ATGTACAGCTGTCTGTAGAGCCGACGAGATGATCCGCTTGGAGCCCACTGGAGGACCC 360	

Db	1681	AGGGGACGGGGGACGGAGACCCCTGCATTTCTCTACGCCACGGTTCCGGAAAGTGTGGA	1740
Qy	1741	GACAGGCCACCGTGCATGACATGGTAGAGAGAGGGGACGGCTTAGGCCCTACACATGCCC	1800
Db	1741	GACAGGGCCACGTCATGACAGTGGAGAGAGGGGCGAGGCCCTGACACACATGCCC	1800
Qy	1801	ACGCTCCCTGCACACGAAAGAGATCCACAACGCTTGGAGAAACACCTCACGTCGTT	1860
Db	1801	ACGCTCCCTGCACACTGAAAGAGATCCACAACCTCTTGGAGAAACACCTCACGTCGTT	1860
Qy	1861	GCGGCACACATTCCTCTCACTCCAGCTCCGCCCATACCCGTCACTACAGCTCACCTCCACCC	1920
Db	1861	GCGGCACACATTCCTCTCACTCCAGCTCCGCCCATACCCGTCACTACAGCTCACCTCCACCC	1920
Qy	1921	CTGTCACATACGGGCTCACCTCCACCCCTGTCACTACAGCTCACCTCCACCTTAA	1980
Db	1921	CTGTCACATACGGGCTCACCTCCACCCCTGTCACTACAGCTCACCTCCACCTTAA	1980
Qy	1981	GTCGCCAGGCCCATGTGCTGGCCTGTCCAAAGGGCTCAAGACTTTTAACTGGAGTGTGTGA	2040
Db	1981	GTCGCCAGGCCCATGTGCTGGCCTGTCCAAAGGGCTCAAGACTTTTAACTGGAGTGTGTGA	2040
Qy	2041	GGGAGTGAAGGTACCTTTGGGGGCAACAGCACCCCTAGTTTCACTCTACCTACCCCTG	2100
Db	2041	GGGAGTGAAGGTACCTTTGGGGGCAACAGCACCCCTAGTTTCACTCTACCTACCCCTG	2100
Qy	2101	CACACATACCTGTGGCAGCGAAATGAAGAAACAGAGCTTCCCGTGCAAAAGGGTCAAGCTC	2160
Db	2101	CACACATACCTGTGGCAGCGAAATGAAGAAACAGAGCTTCCCGTGCAAAAGGGTCAAGCTC	2160
Qy	2161	GCACCCCGCCGCCCTCCCTGCACTCTGCTCCCTCCAGTTCACTTCCGAGCAGCGCA	2220
Db	2161	GCACCCCGCCGCCCTCCCTGCACTCTGCTCCCTCCAGTTCACTTCCGAGCAGCGCA	2220
Qy	2221	GCGCAGGCAACAGTGGCCCAAGAGGCAAGGATCTCTACGGCCCAAGCCGGGGAGG	2280
Db	2221	GCGCAGGCAACAGTGGCCCAAGAGGCAAGGATCTCTACGGCCCAAGCCGGGGAGG	2280
Qy	2281	CTGGAAGGGGTGGCAGATGCGTTCCTCATGCCACGCCAGCGGTCTTGGCTCTCT	2340
Db	2281	CTGGAAGGGGTGGCAGATGCGTTCCTCATGCCACGCCAGCGGTCTTGGCTCTCT	2340
Qy	2341	GTCGCCAGACCTCTGTGACACACAGCCAGATCACAGGGCACCAGGCGAGAGATGCTT	2400
Db	2341	GTCGCCAGACCTCTGTGACACACAGCCAGATCACAGGGCACCAGGCGAGATGCTT	2400
Qy	2401	CTTTTGTCTCTTCTGCGCCGTGGCTGAGCAGATTTTCTTACCCTTAAGATCTGGCTT	2460
Db	2401	CTTTTGTCTCTTCTGCGCCGTGGCTGAGCAGATTTTCTTACCCTTAAGATCTGGCTT	2460
Qy	2461	TGTACTAGAGAAATAAACACATTTTATATTGGTTAAATAAAAAAAAAAAAAAAAAA	2520
Db	2461	TGTACTAGAGAAATAAACACATTTTATATTGGTTAAATAAAAAAAAAAAAAAAAAA	2520
Qy	2521	AAAAAAAAAAAAAAAAAAAAA 2540	
Db	2521	AAAAAAAAAAAAAAAAAAAAA 2540	
RESULT 2			
ABO73249			
ID	ABO73249	standard; cDNA; 2704 BP.	
XX	ABO73249;		
AC	ABO73249;		
XX	30-SEP-2002	(first entry)	
DT			
DE	Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:1.		
KM	Human: phosphatase; mitogen activated protein kinase phosphatase;		
XX	MAP kinase; enzyme; chromosome 11; gene; ss.		

OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	5'UTR
FT	1..93
FT	/tag= a
FT	CDS
FT	94..1509
FT	/tag= b
FT	/product= "MAP kinase phosphatase splice form 1"
FT	1510..2704
FT	3'UTR
FT	/tag= c
PX	
PN	MO200242436-A2.
PD	
XX	30-MAY-2002.
XX	
XX	07-NOV-2001; 2001WO-US42995.
XX	
PR	20-NOV-2000; 2000US-0715177.
PR	18-JAN-2001; 2001US-0761640.
XX	
PA	(PEKE) PE CORP NY.
PI	
PI	WeI M, Ketchum KA, Di Francesco V, Beasley EM;
DR	WPI: 2002-575237/61.
DR	P-PSDB; ABP51653.
XX	
PT	Novel isolated human phosphatase useful for treating disorder
PT	characterized by absence of, inappropriate or unwanted expression of
PT	the phosphatase protein, and as immunogens to raise antibodies -
XX	
PS	Claim 1; Fig 1A; 85dp; English.
XX	
CC	The present invention describes an isolated human phosphatase peptide
CC	(I). (I) can be used for identifying a modulator of (I) by contacting
CC	(I) with an agent and determining if the agent has modulated the
CC	function or activity of (I). (I) is useful for identifying an agent that
CC	binds to (I), by contacting (I) with an agent and assaying the contacted
CC	mixture to determine whether a complex is formed with the agent bound
CC	(I). The human phosphatases from the present invention are mitogen
CC	activated protein (MAP) kinase phosphatases. These human MAP kinase
CC	phosphatases are located on chromosome 11. (I) and the polynucleotide
CC	sequences encoding (I) can be used in gene therapy. The present sequence
CC	encodes human MAP kinase phosphatase splice form 1 from the present
CC	invention.
XX	
SO	Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other:
	Query Match 88.4%; Score 2246.4; DB 24; Length 2704;
	Best Local Similarity 92.4%; Pred. No. 0;
	Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5;
OY	1 CCTGTCCTCGGGGTCCAGACTGCCCGGGGTTGAAGGAAGGGCCCTCCGGTGC 60
Db	CCTGTCCTCGGGGTCCAGACTGC-CGGGGGTTGAAGGAAGGGCCCTCCGGTGC 66
OY	8 CCGTCTCTCGGGGTCCAGACTGC-TCCGGGGGTTGAAGGAAGGGCCCTCCGGTGC 66
OY	61 CAGCCAGAGTGCTCCGGGCTCTGCATATGGGCTCGTACACAGTAGAGCCGTCCGCCCGG 120
Db	CAGCCAGAGTGCTCCGGGCTCTGCATATGGGCTCGTACACAGTAGAGCCGTCCGCCCGG 126
OY	67 CAGCCAGAGTGCTCCGGGCTCTGCATATGGGCTCGTACACAGTAGAGCCGTCCGCCCGG 126
OY	121 GGACAGCGCGCTCCAGGCGCCGTGGGGCCCTGGGACCAGAGCGGCTCCAGGGAAGAGTCGA 180
Db	GGACAGCGCGCTCCAGGCGCCGTGGGGCCCTGGGACCAGAGCGGCTCCAGGGAAGAGTCGA 186
OY	127 GGCACACGCGCGCTCCAGGCGCCGTGGGGCCCTGGGACCAGAGCGGCTCCAGGGAAGAGTCGA 186
OY	181 CTCACGCAAGGCAAGCTTTCGGTCTCCCTGGGGCTGTCTCTGGAGCTGCAGAGATGGA 240
Db	CTCACGCAAGGCAAGCTTTCGGTCTCCCTGGGGCTGTCTCTGGAGCTGCAGAGATGGA 246
OY	187 CTCACGCAAGGCAAGCTTTCGGTCTCCCTGGGGCTGTCTCTGGAGCTGCAGAGATGGA 246
OY	241 GGGGCAATGATGATGACAGAGAGCCCATGTTTGAGCCAACAGAGAAGGCCCGCAGGTGAG 300
Db	GGGGCAATGATGATGACAGAGAGCCCATGTTTGAGCCAACAGAGAAGGCCCGCAGGTGAG 306
OY	247 GGGGCAATGATGATGACAGAGAGCCCATGTTTGAGCCAACAGAGAAGGCCCGCAGGTGAG 306
OY	301 GAGGAGCTTCACGGGGACAGACAGACTTCGGGGCAAGGATCCAGAGTCCCAGAACAG 360

Db 307 GAGAGCTCCACGGGGACGACAGACTTGGGCAAGATCCAGAGTCCCGGAAAGCAG 366
QY 361 GAGAGAGAGGAGGACGACCTGCACTTCATGTATCAGCTTGTAGGGGAGAGATGACATC 420
Db 367 GAGAGAGAGGAGGAGCCTGCACTTCATGTATCAGCTTGTAGGGGAGAGATGACATC 426
QY 421 CGCTTGACAGCCAGCTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTT 480
Db 427 CGCTTGACAGCCAGCTGAGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTT 486
QY 481 TCTACGAGGAGAGAGAGAGTCTGAGCCAGATGAGACGCTCTGAGGCTGAGATTTC 540
Db 487 TCTACGAGGAGAGAGAGAGTCTGAGCCAGATGAGACGCTCTGAGGCTGAGATTTC 546
QY 541 CCTGACAGAGCTCCCGGCTGCAACCTCGGGCTGCTTTCGCCCTCGAGAGTACAC 600
Db 547 CCTGACAGAGCTCCCGGCTGCAACCTCGGGCTGCTTTCGCCCTCGAGAGTACAC 606
QY 601 CAGGTGTATTTAGATGAGAGCGGGGCTTCAGCCTGACGTGTGGTGGCAAGCGGATC 660
Db 607 CAGGTGTATTTAGATGAGAGCGGGGCTTCAGCCTGACGTGTGGTGGCAAGCGGATC 666
QY 661 TTCAAGCCCATCTCCATCCAGACATGT----- 688
Db 667 TTCAAGCCCATCTCCATCCAGACATGTGGGCCACACTCCAGGTATTGACCAAGCATGT 726
QY 689 ----- 688
Db 727 GAGCAGCTTAGGACAGCGGCTTTGTACCGGGGTGAGTGCCCTCAGCTGGCCAGCCAC 786
QY 689 ----- 688
Db 787 TACCAGAGAGACTGACTCCGAACAGAGCTGCCTCAATGATGAGAGGCTATGAGCCGAC 846
QY 689 ----- 688
Db 847 CTGAGTCTGTGGGCTCCAGCGCCGAGCCTGGGGGTCCTCAAGACAGAGCAGATG 711
QY 712 GAGCAGCGATCCGTGTGAGCTGTGAAAGTGTGATGTCACTGACCTGAGAGTGT 771
Db 907 GAGCAGCGATCCGTGTGAGCTGTGAAAGTGTGATGTCACTGACCTGAGAGTGT 771
QY 772 ACTTCCAAAGATCCGCGAGGCTGAGAGTGCCTGGGCTCCCGCTCCAGAGATG 831
Db 967 ACTTCCAAAGATCCGCGAGGCTGAGAGTGCCTGGGCTCCCGCTCCAGAGATG 831
QY 832 CGTACCTTCATGACAAACAGATGCTGCTGTGTGACACAGCGGAGCGAGCCTCCGC 891
Db 1027 CGTACCTTCATGACAAACAGATGCTGCTGTGTGACACAGCGGAGCGAGCCTCCGC 1086
QY 892 ATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGCGAGCAAACTGAGAGAGCTGCAG 951
Db 1087 ATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGCGAGCAAACTGAGAGAGCTGCAG 951
QY 952 AGGACAGAGGTACCCACATCTTGAACATGGCCGGGAGATTGACACTTACCTGAG 1011
Db 1147 AGGACAGAGGTACCCACATCTTGAACATGGCCGGGAGATTGACACTTACCTGAG 1011
QY 1012 CGCTTCACCTACCAAAATGTGGCTTGGGATGAGGATGGGGCCAGCTGCTGCCAC 1071
Db 1207 CGCTTCACCTACCAAAATGTGGCTTGGGATGAGGATGGGGCCAGCTGCTGCCAC 1266
QY 1072 TGGAAAGGAGGACCGCTTATTGAGGCTGCAAGAGCACAGGCGACCCAGCTGTGTC 1131
Db 1267 TGGAAAGGAGGACCGCTTATTGAGGCTGCAAGAGCACAGGCGACCCAGCTGTGTC 1326
QY 1132 CAGTGCAGATGGGCTCAGCGGCTCAGCGGCAACAGTGTGGCTATGCTATGAAGCAG 1191
Db 1327 CAGTGCAGATGGGCTCAGCGGCTCAGCGGCAACAGTGTGGCTATGCTATGAAGCAG 1386
QY 1192 TACGAATGAGGCTGAGAGCGGCTTGGCCAGCTGAGAGGCTCGGGCCATCGGCCCG 1251
+ + + + +

Db 1387 TACGATCCACCTTGAGACAGGCGCCCTGGCCAGCTGACAGAGTCCGCCCATCGCCCG 1446
QY 1252 CCCAACCCCTGCTTCTGCGCAGCTGAGATCTTACAGGCGATCTGACGCGCAGAAC 1311
Db 1447 CCCAACCCCTGCTTCTGCGCAGCTGAGATCTTACAGGCGATCTGACGCGCAGAAC 1506
QY 1312 TGAGGCTGTGGGAGAGAGAGTGTGAGCATGAGAGAGGCGAGCCCGGAAAGA 1371
Db 1507 TGAGGCTGTGGGAGAGAGAGTGTGAGCATGAGAGAGGCGAGCCCGGAAAGA 1566
QY 1372 AGAGCTGGGGCCACGGGCGACCTTAACTTCCGGGGGTCTATGAGTGCATGACTTT 1431
Db 1567 AGAGCTGGG--CCAGCGCCAGCTATTAACCTTCGAGGGGTCTATGAGTGCATGACTTT 1624
QY 1432 CTGAGCGCTTCTTGGAGCTGAGACACCTGAGACCATGAGACCATGCCAGAGTGT 1491
Db 1625 CTGAGCGCTTCTT--GAGGCTGAGAGACACTCAG--AGACAGTGCATGCCAGAGTGT 1682
QY 1492 TCTCTTCCACGAGTCTTTCATGATGAAGAGCCCTGAGCGCTTCCACAGCTTGCAGA 1551
Db 1683 TCTCTTCCACGAGTCTTTCATGATGAAGAGCCCTGAGCGCTTCCACAGCTTGCAGA 1742
QY 1552 CCAAGGAGGCGCAGCAGCTGAGACAGGCGGCTCAGCCTGCTGAAAGTCCCGCAGT 1611
Db 1743 CCAAGGAGGCGCAGCAGCTGAGACAGGCGGCTCAGCCTGCTGAAAGTCCCGCAGT 1802
QY 1612 TGTATTCCCTCCAGGCGAGTGCCTGAGTGGCCCAACCGGACCCAGCGCTTCCAGAGCAG 1671
Db 1803 TGTATTCCCTCCAGGCGAGTGCCTGAGTGGCCCAACCGGACCCAGCGCTTCCAGAGCAG 1862
QY 1672 AGCAGGCGGAGGCGCAGGCGCAGGAGAGGCGCTGATTTCTTACGCGCCAGTTCGGGA 1731
Db 1863 AGCAGGCGGAGGCGCAGGCGCAGGAGAGGCGCTGATTTCTTACGCGCCAGTTCGGGA 1922
QY 1732 AGTGTGAGACAGGCGCAGGCGTGCATGACATGAGAGAGGCGCAGGCGCTGAGCCCTCA 1791
Db 1923 AGTGTGAGACAGGCGCAGGCGTGCATGACATGAGAGAGGCGCAGGCGCTGAGCCCTCA 1982
QY 1792 CACATGCCACGCTCCCTGACACTGAAAGAGATCCCAACTCCTTGGAGAAACCCCTC 1851
Db 1983 CACATGCCACGCTCCCTGACACTGAAAGAGATCCCAACTCCTTGGAGAAACCCCTC 2042
QY 1852 ACGTGTGTCGCGACACTTCCCTGACACTGAGGCGCTCCGCTTACCTGACAGCCTCAC 1911
Db 2043 ACGTGTGTCGCGACACTTCCCTGACACTGAGGCGCTCCGCTTACCTGACAGCCTCAC 2102
QY 1912 CTCCACCCCTGTCTACTACAGGCTCACCTCCACCCCTGTCTACTAGAGCTCACCTCA 1971
Db 2103 CTCCACCCCTGTCTACTACAGGCTCACCTCCACCCCTGTCTACTAGAGCTCACCTCA 2162
QY 1972 CAGCCTTAAGTCCAGGCGCATGTCTGCTGTCTCAAGGCGCTCAAGACTTCTTACCTGGGA 2031
Db 2163 CAGCCTTAAGTCCAGGCGCATGTCTGCTGTCTCAAGGCGCTCAAGACTTCTTACCTGGGA 2222
QY 2032 TGTGTGAGAGGAGCTAAGGTACTTGGGGCAACAGACCTGATGATCTTCACT 2091
Db 2223 TGTGTGAGAGGAGCTAAGGTACTTGGGGCAACAGACCTGATGATCTTCACT 2282
QY 2092 CTGAGCCCTGCACACTACCTGTGTGCGACGAGATGAAGAGAGCTTCCGTCAAAAGAGG 2151
Db 2283 CTGAGCCCTGCACACTACCTGTGTGCGACGAGATGAAGAGAGCTTCCGTCAAAAGAGG 2342
QY 2152 TCAGCGCTTCCACCCCGCCCTCTGCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2211
Db 2343 TCAGCGCTTCCACCCCGCCCTCTGCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2402
QY 2212 AACCGCCACAGGCGACAGCAACAGATGCCCCAAAGAGCAGGAGGATCTCTCAGGCCAGC 2271
Db 2403 AACCGCCACAGGCGACAGCAACAGATGCCCCAAAGAGCAGGAGGATCTCTCAGGCCAGC 2462
QY 2272 CGCGGAGGCTGGAAGGCTGGCAGATGCTTCCCTCATCTCCACCGGTCCAGGTC 2331
Db 2463 CGCGGAGGCTGGAAGGCTGGCAGATGCTTCCCTCATCTCCACCGGTCCAGGTC 2522

Oy	2332	TTTCTGCTGTCCCGACAGCTCCTTGAGACACGCCAGATCACAGGGCCACCAGGCAGA	2391
Dd	2553	TTTTCTCTCTTCCCCAGACTCCTTGAGACACGCCAGATCACAGGGCCACCAGGCAGA	2582
Oy	2392	GATAGTCCTTTTGTGTCCTTTCTGGCCCTCGCTAGTCAGTTTTTCATAGCCTTAGAGT	2451
Dd	2583	GATAGTCCTTTTGTGTCCTTTCTGGCCCTCGCTAGTCAGTTTTTCATAGCCTTAGAGT	2642
Oy	2452	ATCTGGCTTTGTACTGAGAAATAAACACATTTTCATATTTGGTTAAAAA	2511
Dd	2643	ATCTGGCTTTGTACTGAGAAATAAACACATTTTCATATAAAAAAAAAAAAAAAA	2702
Oy	2512	AA 2513	
Dd	2703	AA 2704	
<hr/>			
RESULT 3			
ID	ABO73252	standard; cDNA; 2704 BP.	
XX	ABO73252;		
XX	30-SEP-2002	(first entry)	
DE	Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.		
XX			
KW	Human: phosphatase; mitogen activated protein kinase phosphatase;		
KM	MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;		
KW	SNP; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	5'UTR	1..93	
FT	CDS	/tag= a 94..1509	
FT	variation	/product= "MAP kinase phosphatase splice form 1" replace(577,A) /tag= C	
FT	variation	/standard_name= "single nucleotide polymorphism (SNP)" replace(1451,G) /tag= d	
FT	3'UTR	/standard_name= "single nucleotide polymorphism (SNP)" 1510..2704 /tag= e	
FT	variation	/tag= e replace(2641,A) /tag= f	
FT	variation	/standard_name= "single nucleotide polymorphism (SNP)"	
PX	MO200242436-A2.		
XX			
PD	30-MAY-2002.		
PE	07-NOV-2001; 2001MO-US42995.		
PR	20-NOV-2000; 2000US-0715177.		
PR	18-JAN-2001; 2001US-0761640.		
PA	(PEKE) PE CORP NY.		
PI	WeI MeI, Ketchum KA, Di Francesco V, Beasley EM;		
DR	WPI: 2002-575237/61.		
DR	P-PsDB: ABP51653.		
XX	Novel isolated human phosphatase peptide useful for treating disorder		
PT	characterized by absence of, inappropriate or unwanted expression of		
PT	the phosphatase protein, and as immunogens to raise antibodies -		
BS	Claim 1: Fig 3a; 85pp; English.		

xx	The present invention describes an isolated human phosphatase peptide
cc	(I). (I) can be used for identifying a modulator of (I) by contacting
cc	(I) with an agent and determining if the agent has modulated the
cc	function or activity of (I). (I) is useful for identifying an agent that
cc	binds to (I), by contacting (I) with an agent and assaying the contacted
cc	mixture to determine whether a complex is formed with the agent bound
cc	(I). The human phosphatases from the present invention are mitogen
cc	activated protein (MAP) kinase phosphatases. These human MAP kinase
cc	phosphatases are located on chromosome 11. (I) and the polynucleotide
cc	sequences encoding (I) can be used in gene therapy. The present sequence
cc	encodes human MAP kinase phosphatase splice form 1 from the present
cc	invention.
xx	
sq	Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other;
	Query Match 88.4%; Score 2246.4; DB 24; Length 2704.
	Best Local Similarity 92.4%; Pred. No. 0;
	Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5
Oy	1 CCTGTCCTCGCGGGTCCAGACTGTCCCACGGGGTTGAGGAAGAAGGCCTGCCGTGC 60
Dd	8 CCTGTCTCCTGGGGTGTCAGACTGT-CCGGGGGTTGAGGAAGAAGGCCTGCCGTGC 66
Oy	61 CAGCCCAAGTGTCTCGCGGCTGGCTCCATATGACCCTTGCTCACAGTAGACCCGTTGCCCCCG 120
Dd	67 CAGCCCAAGTGTCTCGCGGCTGGCTCCATATGACCCTTGCTCACAGTAGACCCGTTGCCCCCG 126
Oy	121 GGCACGGGGGCTCCACAGCGCCGGGGGGCCCTGGGACACAGGGGGTCCAGCGAAGAGATGCA 180
Dd	127 GGCACGGGGGCTCCACAGCGCCGGGGGGCCCTGGGACACAGGGGGTCCAGCGAAGAGATGCA 186
Oy	181 CTCACAAGCAAGCAGAGCTTTTCCGGTGCTCCCTGGGGCTGTCTCTGGACATGCAGATGGA 240
Dd	187 CTCACAAGCAAGCAGAGCTTTTCCGGTGCTCCCGGGGCTGTCTCTGGACATGCAGATGGA 246
Oy	241 GGGGCAATGATGATGCACAGAGGCAGTTCGTGAGCCACAAGAGAAGGCCCCGAGTGG 300
Dd	247 GGGGCAAAATGATGATGCACAGAGGCAGCTTTGTGAGCCAAGAGAAGGCCCCGAGTGG 306
Oy	301 GAGGAGCTTCACGGGGACACAGACAGACTTCCGGCAAGAGATCCACAGAGTCCACAGAAGCAG 360
Dd	307 GAGGAGCTTCACGGGGACACAGACAGACTTCCGGCAAGAGATCCACAGAGTCCACAGAAGCAG 366
Oy	361 GAGGAGCAGAGGACAGCACTGCACTCATGTATGTAACAGCTCTGAGGCCGAGAGATGATC 420
Dd	367 GAGGAGCAGAGGACAGCACTGCACTCATGTATGTAACAGCTCTGAGGCCGAGAGATGATC 426
Oy	421 GCCCTGGGCAAGCAGTGGAGGACACCCGGGCTCCCCGGCTCCGCTACTCGTGAGATT 480
Dd	427 GCCCTGGGCAAGCAGTGGAGGACACCCGGGCTCCCCGGCTCCGCTACTCGTGAGATT 486
Oy	481 TCATACGAGAAGAGAAAGTCTGAGCCAGATGAGACAGTCCCTCTGGGCTGTGATTT 540
Dd	487 TCATACGAGAAGAGAAAGTCTGAGCCAGATGAGACAGTCCCTCTGGGCTGTGATTT 546
Oy	541 CCTGACAGAGCTCCCCCAAGCTGACACCTTGGGCTGTGTTTGCCCTCTTGAGATGACACC 600
Dd	547 CCTGACAGAGCTCCCCCAAGCTGACACCTTGGGCTGTGTTTGCCCTCTTGAGATGACACC 606
Oy	601 CAGGTACTTATGATGAGAACGGGGGCTTCACAGTACGCTTGATGGGCAAAAGCCGGATC 660
Dd	607 CAGGTACTTATGATGAGAACGGGGGCTTCACAGTACGCTTGATGGGCAAAAGCCGGATC 666
Oy	661 TTCAGAGCCATCTCATTCAGACCATGT----- 688
Dd	667 TTCAGAGCCATCTCATTCAGACCATGTGGCCACACTCCAGGTATTTCACCAAGCATGT 726
Oy	689 ----- 688
Dd	727 GAGGCAAGCTCTAAGCAGCGGCTTGTACCGGGGTGGCATGTGCCCTCACTCGGGCCAGCCAC 786
Oy	689 ----- 688

Db 787 TACCAGAGAGACTGAACCTCCGAACAGAGCTGGCTCAATGATGAGAGGCTATGGCCGAC 816
QY 689 -----GCTCTCAGAAACAGGAGCATG 711
Db 847 CTGGAGTCTTCGGGCGCTCCAGCGCCGAGCCCTGGCGGGGCTCTCAGAAACAGGAGCATG 906
QY 712 GAGCAGGCGATCCGCTGCTGAGCTGTGAAAGTGTGGATGTCAAGTCAACCTGGAGAGTGTG 771
Db 907 GAGCAGGCGATCCGCTGCTGAGCTGTGAAAGTGTGGATGTCAAGTCAACCTGGAGAGTGTG 966
QY 772 ACTTCCAAAGAGATCCGCCAGAGCTGTGAGAGTGTGGAGTGTGGAGTGTGGAGTGTG 831
Db 967 ACTTCCAAAGAGATCCGCCAGAGCTGTGAGAGTGTGGAGTGTGGAGTGTGGAGTGTG 1026
QY 832 CGTGAATTCATCAGCAACAGATGTGCTGTGCTGTGGAGAGTGTGGAGAGAGGAGAGTGTG 891
Db 1027 CGTGAATTCATCAGCAACAGATGTGCTGTGCTGTGGAGAGTGTGGAGAGAGGAGAGTGTG 1086
QY 892 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAACGACGAACAACTGGAGAGTGTGAG 951
Db 1087 ATCTTCCCGCCACCTTACCTGGGCTCAGAGTGAACGACGAACAACTGGAGAGTGTGAG 1146
QY 952 AGGAACAGGAGTACCCACATCTTTGAACATGTGCCCGGAGATTTGACAACTTCAACCTGAG 1011
Db 1147 AGGAACAGGAGTACCCACATCTTTGAACATGTGCCCGGAGATTTGACAACTTCAACCTGAG 1206
QY 1012 CGCTTACCTACCAACATGTGCGCTCTGGATGAGAGTGTGGAGAGTGTGGAGAGTGTG 1071
Db 1207 CGCTTACCTACCAACATGTGCGCTCTGGATGAGAGTGTGGAGAGTGTGGAGAGTGTG 1266
QY 1072 TGGAGAGAGAGCAGCCGCTTCAATTAGAGCTGTGAAGACACAGAGGACCCACGCTGTGTC 1131
Db 1267 TGGAGAGAGAGCAGCCGCTTCAATTAGAGCTGTGAAGACACAGAGGACCCACGCTGTGTC 1326
QY 1132 CACTCAGAGATGGGGGTGAGCGCTCAGCGGCGCACAGTGTGGCTATGGCATGAGAGAG 1191
Db 1327 CACTCAGAGATGGGGGTGAGCGCTCAGCGGCGCACAGTGTGGCTATGGCATGAGAGAG 1386
QY 1192 TACGATATGAGCTGTGAGCAGAGGCTGTGGAGAGTGTGGAGAGTGTGGAGAGTGTG 1251
Db 1387 TACGATATGAGCTGTGAGCAGAGGCTGTGGAGAGTGTGGAGAGTGTGGAGAGTGTG 1446
QY 1252 CCAACCTTGGGCTTCTGCGGCGAGTGTGAGATCTTCAACAGAGGCTATGAGCGCCAGAAC 1311
Db 1447 CCAACCTTGGGCTTCTGCGGCGAGTGTGAGATCTTCAACAGAGGCTATGAGCGCCAGAAC 1506
QY 1312 TGAGGGTGTGGGAGAGAAAGTGTGAGCATGGAAGAGAGAGCAGAGCAGCCGAAAGA 1371
Db 1507 TGAGGGTGTGGGAGAGAAAGTGTGAGCATGGAAGAGAGAGCAGAGCAGCCGAAAGA 1566
QY 1372 AGAGCTGGGGGCGACAGGGGCGACGTATAACCTCCGAGGGGTATGAGGTGCATAGTCTT 1431
Db 1567 AGAGCTGGGGGCGACAGGGGCGACGTATAACCTCCGAGGGGTATGAGGTGCATAGTCTT 1624
QY 1432 CTGGAGCCTTCTTGGAGCTGTGAGACACCTCAGTACAGAGTGTGAGATGAGAGTGTG 1491
Db 1625 CTGGAGCCTTCTTGGAGCTGTGAGACACCTCAGTACAGAGTGTGAGATGAGAGTGTG 1682
QY 1492 TCTTCTCCACGAGTCTTCAATGAAAGAGCTGTGAGAGCCTTCCACAGTTTGCAAGA 1551
Db 1683 TCTTCTCCACGAGTCTTCAATGAAAGAGCTGTGAGAGCCTTCCACAGTTTGCAAGA 1742
QY 1552 CCAAGGAGGCGACAGAGTGTGAGAGAGGGGCTCAGAGCTGTGAGAGTGTGAGAGTGTG 1611
Db 1743 CCAAGGAGGCGACAGAGTGTGAGAGAGGGGCTCAGAGCTGTGAGAGTGTGAGAGTGTG 1802
QY 1612 TGGTTACCTTCAAGGAGTGTGAGTGTGAGAGAGGAGAGGAGAGGAGAGGAGAGG 1671
Db 1803 TGGTTACCTTCAAGGAGTGTGAGTGTGAGAGAGGAGAGGAGAGGAGAGGAGAGG 1862
QY 1672 AGCAGGGGAGAGGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGG 1731
|||||

Db 1863 AGCAGGGGAGAGGGGAGAGGGGAGAGAGCCCTGCATTTCTCTACGCCAGGTTCCGGA 1922
QY 1732 AGTGTGTGAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGAGGCGAGGCTTACCCCTTA 1791
Db 1923 AGTGTGTGAGACAGGCGCAGCGTGCATGACAGTGGAGAGAGGAGGCGAGGCTTACCCCTTA 1982
QY 1792 CACATCCCAAGCTCCCTCCAGACAGTGAAGAGATCCACAACTCTTGGAGAAACACCTTC 1851
Db 1983 CACATCCCAAGCTCCCTCCAGACAGTGAAGAGATCCACAACTCTTGGAGAAACACCTTC 2042
QY 1852 AGCTGTGGCGGACACATTTCTCTCAGCTCCGCCCCATACCCCTCAGTACAGCTTAC 1911
Db 2043 AGCTGTGGCGGACACATTTCTCTCAGCTCCGCCCCATACCCCTCAGTACAGCTTAC 2102
QY 1912 CTCCACCCCTTGTCACTACAGGCTCAGCTCCACCCCTGTGCTACAGCTTACCTCTTA 1971
Db 2103 CTCCACCCCTTGTCACTACAGGCTCAGCTCCACCCCTGTGCTACAGCTTACCTCTTA 2162
QY 1972 CAGCTTAAAGTCCCAAGGCGCATGTGCTGCTGCAAGGAGTCAAGACTTTCACTGAGGA 2031
Db 2163 CAGCTTAAAGTCCCAAGGCGCATGTGCTGCTGCAAGGAGTCAAGACTTTCACTGAGGA 2222
QY 2032 TGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCCTAGTTTCAATTCTCAACT 2091
Db 2223 TGTGTGAGAGGAGTGAAGTACCTTTGGGGGCAACAGCACCCTAGTTTCAATTCTCAACT 2282
QY 2092 CTAGCCTTGCACACTCAGCTGTGGCAGCGAATGAACAGAGCTTCCGTCGCAAAAAGG 2151
Db 2283 CTAGCCTTGCACACTCAGCTGTGGCAGCGAATGAACAGAGCTTCCGTCGCAAAAAGG 2342
QY 2152 TCAGGCTTCCACCCCGGCGGCTCCCTGCGACCTCTGCTCTCTCCAGTTCATTTCTG 2211
Db 2343 TCAGGCTTCCACCCCGGCGGCTCCCTGCGACCTCTGCTCTCTCCAGTTCATTTCTG 2402
QY 2212 AACCAAGCAGGCGCAGGCAACCAAGTGGCCCCCAAGAGGAGAGAGATCCTAGAGCCGAGC 2271
Db 2403 AACCAAGCAGGCGCAGGCAACCAAGTGGCCCCCAAGAGGAGAGATCCTAGAGCCGAGC 2462
QY 2272 GCGGGAGGGGTGGAAGGCTGTGGAGATGCTTCTCTCTATCCACTCCACGGGTCAGAGTC 2331
Db 2463 GCGGGAGGGGTGGAAGGCTGTGGAGATGCTTCTCTCTATCCACTCCACGGGTCAGAGTC 2522
QY 2332 TTTGCTGCTTCCACAGACCTCTGTGACACACAGCCAGATCAGAGGCGACAGGCGCAGA 2391
Db 2523 TTTGCTGCTTCCACAGACCTCTGTGACACACAGCCAGATCAGAGGCGACAGGCGCAGA 2582
QY 2392 GATAGTCTTCTTTTGTGCTTTTGTGCGCTGTGGCTAGTCAAGTTCATAGCCTTACAGT 2451
Db 2583 GATAGTCTTCTTTTGTGCTTTTGTGCGCTGTGGCTAGTCAAGTTCATAGCCTTACAGT 2642
QY 2452 ATCTGCTTTGTACTGAGAAATAAACATTTTCATATTTGGTTAAAAAAGAAAAA 2511
Db 2643 ATCTGCTTTGTACTGAGAAATAAACATTTTCATATTTGGTTAAAAAAGAAAAA 2702
QY 2512 AA 2513
Db 2703 AA 2704
RESULT 4
AAD36063
ID AAD36063 standard; cDNA; 2618 BP.
AC AAD36063;
XX
DT 09-AUG-2002 (first entry)
XX
DE Murine dual-specificity phosphatase 15 (Dsp-15) cDNA.
XX
KW Murine: dual-specificity phosphatase 15; DSP15; antiallergic; cytosolic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;

KW metabolic disease; allergy; screening; gene; ss.
XX Mus musculus.
OS
FH Key Location/Qualifiers
FT CDS 35..1450
FT /*tag= a
FT /product= "Murine DSP-15 protein"
XX W0200224740-A2.
XX
XX 28-MAR-2002.
XX
XX 19-SEP-2001; 2001WO-US29406.
XX
XX 19-SEP-2000; 2000US-233833P.
XX 18-SEP-2001; 2001US-0955732.
XX
XX (CEPT-) CEPTYR INC.
XX
XX Luche RM, Wei B;
XX
XX WPI: 2002-394127/42.
XX P-PSDB: AAE22733.
XX
XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
XX PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
XX PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
XX PT diseases
XX
XX Claim 56; Fig 4; 91pp; English.
XX
XX The invention relates to a new isolated dual-specificity phosphatase 15
XX (DSP-15) polypeptide which retains the ability to dephosphorylate an
XX activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
XX that dephosphorylate both phosphotyrosine and phosphothreonine/serine
XX residues. DSP-15 polypeptides may be used to identify agents that
XX modulate DSP-15 activity, where such agents may inhibit or enhance signal
XX transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
XX polypeptides, modulating agents, and/or polynucleotides encoding the
XX polypeptides may be used to modulate DSP-15 activity in a patient, and to
XX ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
XX versus-host disease, autoimmune diseases, allergies, metabolic diseases,
XX abnormal cell growth, abnormal cell proliferation and cell cycle
XX abnormalities. DSP-15 alternate form polypeptides are useful in screening
XX assays for modulators of enzyme activity and/or substrate binding. The
XX present sequence is murine DSP-15 cDNA.
XX
XX Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other;
XX
XX
XX Query Match 86.2%; Score 2189.6; DB 24; Length 2618;
XX Best Local Similarity 92.5%; Pred. No. 0;
XX Matches 2425; Conservative 0; Mismatches 4; Indels 193; Gaps 4;
XX
XX
XX 54 CCGGTGCCAGCCAGGTGCTCGCGGCTGCTCCATGGCCCTGTGCACATGAGCCGTTTC 113
XX 1 CCGGTGCCAGCCAGGTGCTCGCGGCTGCTCCATGGCCCTGTGCACATGAGCCGTTTC 60
XX
XX 114 GCCCCCGGCGAGCGGCTCCACGCGCGTGGGGCCCTGGGACCAAGCGGCTCCAGCAAG 173
XX 61 GCCCCCGGCGAGCGGCTCCACGCGCGTGGGGCCCTGGGACCAAGCGGCTCCAGCAAG 120
XX
XX 174 GAGTGCATCTCAGGAGGAGGAGGCTTGGGCTGCTCCGTGGGGCTGTCCTGGGAGCTCA 233
XX 121 GAGTGCATCTCAGGAGGAGGAGGCTTGGGCTGCTCCGTGGGGCTGTCCTGGGAGCTCA 180
XX
XX 234 GGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 293
XX 181 GGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
XX
XX 294 GAGTGCATCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 353
XX 241 GAGTGCATCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 354 GAAGCAGGAGAGCAGAGCAGACCTGACCTCATGTGATACAGCTGTGAGGCCAGAGA 413
DB 301 GAAGCAGGAGAGCAGAGCAGACCTGACCTCATGTGATACAGCTGTGAGGCCAGAGA 360
QY 414 TGACATCCGCTGGGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTTACTGCT 473
DB 361 TGACATCCGCTGGGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTTACTGCT 420
QY 474 GGTATTTCTACACAGAGAGAGAGAGTGTGAGCCAGGATGAGAGGTCTCTCTGGGGCT 533
DB 421 GGTATTTCTACACAGAGAGAGAGAGTGTGAGCCAGGATGAGAGGTCTCTCTGGGGCT 480
QY 534 GGATTTCCCTACACAGACCTCCCGGCTGGGAGCCCTGGGCTGGCTGGCCCTCTGGAG 593
DB 481 GGATTTCCCTACACAGACCTCCCGGCTGGGAGCCCTGGGCTGGCTGGCCCTCTGGAG 540
QY 594 TGACACCCAGGAGTACTAGATGAGAGGAGGAGGCTTACGCTGACGCTGTGGGCAAG 653
DB 541 TGACACCCAGGAGTACTAGATGAGAGGAGGAGGCTTACGCTGACGCTGTGGGCAAG 600
QY 654 CCGGATCTTCAAGCCCATCTCCATCCAGCATG----- 688
DB 601 CCGGATCTTCAAGCCCATCTCCATCCAGCATG----- 688
QY 689 ----- 688
DB 601 AGCATGTAGGAGCAGCTTAGGACAGGCGCTTGTACCGGGTGGAGTGCCTTACCTGGGC 720
QY 689 ----- 688
DB 721 CAGCCACTACAGAGAGAGATGAACTCCGAACAGAGTGTCTCAATGATGAGGAGCTAT 780
QY 689 ----- 688
DB 781 GAGCCAGCTGAGTCTGCGGCTCCAGGCGGAGCCGAGCTGGCGGCTCTAGAAACAGA 840
QY 705 GAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764
DB 841 GAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 765 GAGTGTCACTTCCAAAGAGATCCGAGGCTGTGGAGCTGGGCGCTCCGCTCCCA 824
DB 901 GAGTGTCACTTCCAAAGAGATCCGAGGCTGTGGAGCTGGGCGCTCCGCTCCCA 960
QY 825 GCAGTACCGTACTTCAATGACACCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 884
DB 961 GCAGTACCGTACTTCAATGACACCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 885 CTCCGCACTTCCGCGGCTGCTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 944
DB 1021 CTCCGCACTTCCGCGGCTGCTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1080
QY 945 GCTGCAGAGAGAGAGGAGTCCACCATCTTGAACATGAGCCCGGAGATGACAACTTCTA 1004
DB 1081 GCTGCAGAGAGAGAGGAGTCCACCATCTTGAACATGAGCCCGGAGATGACAACTTCTA 1140
QY 1005 CCTGTAGGCTTTCACCTTACCAATGTGCGCTCTGGGATGAGAGTGGGCGCACTGCT 1064
DB 1141 CCTGTAGGCTTTCACCTTACCAATGTGCGCTCTGGGATGAGAGTGGGCGCACTGCT 1200
QY 1065 GCGGCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1124
DB 1201 GCGGCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1125 GCTGTCCACTGCAAGATGGGCTGAGCGCTCAGCGGCTCAGAGTGTGGCTTATGCTAT 1184
DB 1261 GCTGTCCACTGCAAGATGGGCTGAGCGCTCAGCGGCTCAGAGTGTGGCTTATGCTAT 1320
QY 1185 GAAGCAGTACCAATGACGCTGAGAGGAGGCGCTGGCGCCAGCTGCAAGAGGCTTCGGGCCAT 1244
DB 1321 GAAGCAGTACCAATGACGCTGAGAGGAGGCGCTGGCGCCAGCTGCAAGAGGCTTCGGGCCAT 1380


```

Db 2098 ACCCTCAGCTGTGTCGCGACACATTCCTCAGCTCGCCCAATACCCGTACACAG 2157
QY 1906 CCTCAGCTCCACCCCTGTACAGGCGCTCAGCTCCACCCCTGTACACAGCTCAG 1965
Db 2158 CCTCAGCTCCACCCCTGTACAGGCGCTCAGCTCCACCCCTGTACACAGCTCAG 2217
QY 1966 CTCTCAGACCTTTAACTCCAGGCGCATGTCTGCTGTCCAGGCGCTCAAGCTTCA 2025
Db 2218 CTCTCAGACCTTTAACTCCAGGCGCATGTCTGCTGTCCAGGCGCTCAAGCTTCA 2277
QY 2026 CTGGATGTGTAGAGGCGCTAAGTACCTTTGGGGCAACAGCACCTAGTTTCATTC 2085
Db 2278 CTGGATGTGTAGAGGCGCTAAGTACCTTTGGGGCAACAGCACCTAGTTTCATTC 2337
QY 2086 TCAACTCTAGCCCTCAGACCTCAGCTGTGCAAGCAATGAAACAGACCTCCGTCGA 2145
Db 2338 TCAACTCTAGCCCTCAGACCTCAGCTGTGCAAGCAATGAAACAGACCTCCGTCGA 2397
QY 2146 AAAGGTCAGGCTCCGACCCCGCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2205
Db 2398 AAAGGTCAGGCTCCGACCCCGCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2457
QY 2206 TCCTGGAACCAAGCCAGGCGCAACAGTGGCCCAAGGCGAGAGTCTCTAGGC 2265
Db 2458 TCCTGGAACCAAGCCAGGCGCAACAGTGGCCCAAGGCGAGAGTCTCTAGGC 2517
QY 2266 CCCAGCGCGGAGGCGTGGGAAGGCTGGCAGATGCTTCCCTCAGCTCCACCGGTC 2325
Db 2518 CCCAGCGCGGAGGCGTGGGAAGGCTGGCAGATGCTTCCCTCAGCTCCACCGGTC 2577
QY 2326 CAGTCTTTCTGCTGTCGCCAGACCTCCTGTACACACAGCAAGTACAGGCGACAG 2385
Db 2578 CAGTCTTTCTGCTGTCGCCAGACCTCCTGTACACACAGCAAGTACAGGCGACAG 2637
QY 2386 GCCAGATATGCTCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTC 2445
Db 2638 GCCAGATATGCTCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTCCTTTTTC 2697
QY 2446 TACAGTATCTGCTTTGTACTGAAATATAACATTTTCATTTTGGTTAAAAAAA 2505
Db 2698 TACAGTATCTGCTTTGTACTGAAATATAACATTTTCATTTTAAAAAAA 2757
QY 2506 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540
Db 2758 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2792

```

```

RESULT 6
AAF30485
ID AAF30485 standard; cDNA; 2781 BP.
XX
AC AAF30485;
XX

```

```

DT 29-MAY-2001 (first entry)
XX

```

```

DE Human protein phosphatase and kinase protein-10 cDNA 5039718CB1.
XX

```

```

KW Protein phosphatase and kinase protein; PPHK-10; human;
KW gastrointestinal disorder; immune system disorder;
KW neurological disorder; cell proliferative disorder; cancer;
KW diagnosis; therapy; ss.
XX

```

```

OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT CDS 87..2066
FT sig_peptide /*lag= a
FT mat_peptide /*lag= b
FT misc_feature /*lag= c
FT /*lag= d

```

```

FT FT /note= "unique fragment"
FT FT 921..980
FT FT /*lag= e
FT FT /note= "unique fragment"
XX
XX WO200120004-A2.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000; 2000WO-US25515.
XX
XX 15-SEP-1999; 99US-0154141.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzal Y;
XX Lu DAM;
XX WPI; 2001-244811/25.
XX P-PSDB; AAB20331.
XX
XX Novel human protein phosphatase and kinase proteins for diagnosis,
XX treatment and prevention of gastrointestinal, immune system,
XX neurological and cell proliferative disorders -
XX
XX Claim 5; Page 101-102; 103pp; English.
XX
XX The present sequence is that of cDNA encoding novel human
XX protein phosphatase and kinase protein PPHK-10 (see AAB20331).
XX The cDNA was initially identified in Incyte Clone ID No. 5039718CB1,
XX from a colon tumour tissue library. Tissues that express PPHK-10
XX (as a fraction of total tissues expressing PPHK-10) include
XX reproductive (0.343), gastrointestinal (0.194) and haematopoietic
XX or immune (0.134). Diseases or conditions associated with tissues
XX expressing PPHK-10 (as a fraction of total tissues expressing
XX cell proliferation (0.090). The encoded protein shows homology to
XX Drosophila melanogaster MAP kinase phosphatase. The invention
XX provides human PPHK-1 to -11 polypeptides (see AAB20332-32) and
XX polynucleotides (see AAF30476-86). It also provides expression
XX vectors, host cells, antibodies, agonists and antagonists, as well
XX as methods for diagnosing, treating or preventing disorders
XX associated with expression of PPHK, including gastrointestinal
XX disorders, immune system disorders, neurological disorders and cell
XX proliferative disorders, including cancer.
XX
XX Sequence 2781 BP; 576 A; 906 C; 820 G; 479 T; 0 other;
XX

```

```

Query Match 83.3%; Score 2115; DB 22; Length 2781;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 5; Indels 292; Gaps 6;

```

```

QY 10 GCGGTCAGAGTCTCCCGGGGCTTGAAGGAGGCGCTGCCGTCAGCCAGG 69
Db 10 GCGGTCAGAGTCTT-CCGCGGGCTTGAAGGAGGCGCTGCCGTCAGCCAGG 68
QY 70 TGTCTCGGCTTGCCTCCTGCTCAATGAGCTGACAGTGGCGCTTCCCGGCGAGCGGC 129
Db 69 TGTCTCGGCTTGCCTCCTGCTCAATGAGCTGACAGTGGCGCTTCCCGGCGAGCGGC 128
QY 130 GCCTCAGGCGCGTGGGCGCTTGGAGCCAGGCGGCTCCAGCAAGAGAGTGCATCCAGCA 189
Db 129 GCCTCAGGCGCGTGGGCGCTTGGAGCCAGGCGGCTCCAGCAAGAGAGTGCATCCAGCA 188
QY 190 AGGCAAGCTTGGCGTCTCCGTGGGCTGCTGCGAGCTGAGATGAGAGGAGCAAT 249
Db 189 AGGCAAGCTTGGCGTCTCCGTGGGCTGCTGCGAGCTGAGATGAGAGGAGCAAT 248
QY 250 GATGATGAGCAGAGGCGCAATTTGAGCCAAAGAGGCGCGAGTGTAGAGAGAGCTC 309
Db 249 GATGATGAGCAGAGGCGCAATTTGAGCCAAAGAGGCGCGAGTGTAGAGAGAGCTC 308
QY 310 CACGGGACCAAGAGACTTCCGGCAAGAGATCCAGATCCCAAGAAAGCAGAGACAG 369

```

|||||
Db 309 CACGGGACACAGAGACTTGGGCGAGATCCAGAGTCCCAAGAAACAGAGAGAGCAG 368
OY 370 AGGAGACACCTGCACCTATGTAAGCTGCTGAGGCGCGAGAGATGATCCGCTGGCA 429
Db 369 AGGAGACACCTGCACCTATGTAAGCTGCTGAGGCGCGAGAGATGATCCGCTGGCA 428
OY 430 GCCAGCTGAGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGATGTTTCTACAGCA 489
Db 429 GCCAGCTGAGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGATGTTTCTACAGCA 488
OY 490 GAAGAGAGAGCTTGAGGCGAGATGAGAGGTCCTCTGGGCGTGATTTCCCTGACAGC 549
Db 489 GAAGAGAGAGCTTGAGGCGAGATGAGAGGTCCTCTGGGCGTGATTTCCCTGACAGC 548
OY 550 AGCTCCCCAGCTGCACCTTGGGCTGCTTTGGCCCTCTGAGAGTGAACCCAGGTGTAC 609
Db 549 AGCTCCCCAGCTGCACCTTGGGCTGCTTTGGCCCTCTGAGAGTGAACCCAGGTGTAC 608
OY 610 TTAGATGAGAGCGGGGCTTGAGCTGACGTCCTGGTGGGCAAGCGGATCTTCAAGGCC 669
Db 609 TTAGATGAGAGCGGGGCTTGAGCTGACGTCCTGGTGGGCAAGCGGATCTTCAAGGCC 668
OY 670 ATCTCCATCCAGAGCCATGT----- 688
Db 669 ATCTCCATCCAGAGCCATGTGGGCGCACACTCCAGTATTGCAACAGCATGTGAGGACACT 728
OY 689 ----- 688
Db 729 CTAGGAGACGGCTGTGTACCGGCTGGAGTGAGTCCCTACCTGGGCGACACTACAGGAG 788
OY 689 ----- 688
Db 789 AGACTGAACCTCCGAACAGAGCTGCTCAATGAGTGAAGGCTATGCGCGACCTGAGACTCT 848
OY 689 -----GTCCTTCAGAACAGAGAGCAGATGGAGAGCG 720
Db 849 CTGCGGCTCCAGTAGCGGAGCGCTGGCGGTCTCAGAACAGAGACAGATGGAGAGCG 908
OY 721 ATCCCTGCTGAGCTGTGGAAGGTGTTGATGTGCTAGTGAACCTGGAGAGTGTCACTTCCAAA 780
Db 909 ATCCCTGCTGAGCTGTGGAAGGTGTTGATGTGCTAGTGAACCTGGAGAGTGTCACTTCCAAA 968
OY 781 GAGATCCGCGAGGCTCTGAGAGCTGGGCTGGGCTCCCTCCAGACAGTACCTGACTTTC 840
Db 969 GAGATCCGCGAGGCTCTGAGAGCTGGGCTGGGCTCCCTCCAGACAGTACCTGACTTTC 1028
OY 841 ATGCAACAACAGATGCTGCTGTGTGGACAGCGGAGCGAGCTCCCGCATTTTCCC 900
Db 1029 ATGCAACAACAGATGCTGCTGTGTGGACAGCGGAGCGAGCTCCCGCATTTTCCC 1088
OY 901 CACCTCTGACCTGGGCTGAGAGTGAAGCAGCAACCTGGAGAGAGCTGACAGAGAAAG 960
Db 1089 CACCTCTGACCTGGGCTGAGAGTGAAGCAGCAACCTGGAGAGAGCTGACAGAGAAAG 1148
OY 961 GTACCCACATCTTGAATATGAGCGCGGAGATGTACAATCTTCACTGGAGGCTTCAAC 1020
Db 1149 GTACCCACATCTTGAATATGAGCGCGGAGATGTACAATCTTCACTGGAGGCTTCAAC 1208
OY 1021 TACCACAATGTGCGGCTGTGGAGTGAAGTGGCCAGCTGCTGCGCACACTGGAAGAG 1080
Db 1209 TACCACAATGTGCGGCTGTGGAGTGAAGTGGCCAGCTGCTGCGCACACTGGAAGAG 1268
OY 1081 ACGCACCGCTTATGAGGCTGCAAGAGCAGAGGCAACCGAGCTGCTGCTGCTACAGAG 1140
Db 1269 ACGCACCGCTTATGAGGCTGCAAGAGCAGAGGCAACCGAGCTGCTGCTGCTACAGAG 1328
OY 1141 ATGGGCTGAGCGGCTGAGCGGCGACAGTGTGCTGATGATGAAGCAGTGAAGTGC 1200
Db 1329 ATGGGCTGAGCGGCTGAGCGGCGACAGTGTGCTGATGATGAAGCAGTGAAGTGC 1388
OY 1201 AGCCTGAGAGAGCGGCTGCGGCGACGTGAGAGGCTCGGCGCCATGCGCGCCCAACCT 1260
|||||

Db 1389 AGCCTGAGAGAGCGGCTTGCAGGCTGAGAGAGTCCGGCCCATGCGCCGCCCAACCT 1448
OY 1261 GGCCTTCCTGGCGCAGCTGACATGTACAGAGGCAATCCGAGC----- 1302
Db 1449 GGCCTTCCTGGCGCAGCTGACATGTACAGAGGCAATCCGAGGCGCAGGCCAGACCAT 1508
OY 1303 ----- 1302
Db 1509 GTCTGGAGAGAAAGTGGGTGGGTCTCCCAAGAGAGCAACCCAGCCCTGAAGTCTCT 1568
OY 1303 -----GCCAGACCTGAGGCTGTGGGAGAGAGAGTGTAGGC 1342
Db 1569 ACACCATTCACACTCTTCCCGCAGAACCTGAGGGTGTGGGAGAGAAAGTTGTAGGC 1628
OY 1343 ATGAGAGAGAGCGAGGCGAGCCCGCAAGAGCCTGGGGGCACCGGGCAGCTATAAAC 1402
Db 1629 ATGAGAGAGAGCGAGGCGAGCCCGCAAGAGAGCCTGGG--CCACGGCGACGTATAAAC 1686
OY 1403 TCCGAGGGGTATGAGGTCATCAAGTCTTGTGAGAGCCCTCTGGAGAGCTGAGAGCACC 1462
Db 1687 TCCGAGGGGTATGAGGTCATCAAGTCTTGTGAGAGCCCTCT--GAGACTGAGAGCACC 1745
OY 1463 TCAGTAGACAGTGAACATGACAGAGGTCCTTCCACGAGATCTTCATGAAGAGCC 1522
Db 1746 TCAG--AGACAGTGAACATGACAGAGGTCCTTCCACGAGTCTTCATGAAGAGCC 1804
OY 1523 TCTGAGCGCTTCCACAGCTTTCAGAGGACCAAGGAGGCCAGCAGAGTGGAGGGGCC 1582
Db 1805 TCTGAGCGCTTCCACAGCTTTCAGAGGACCAAGGAGGCCAGCAGAGTGGAGGGGCC 1864
OY 1583 TCAGCTGCCCTGAAAGTCCCGCAGTCAAGTGTACCTTCAGGCGCAGTGCCTGTGGC 1642
Db 1865 TCAGCTGCCCTGAAAGTCCCGCAGTCAAGTGTGTACCTTCAGGCGCAGTGCCTGTGGC 1924
OY 1643 CAACCGAGCCAGGCTTCCAGAGACAGAGGAGGAGGCGCAGGAGAGAGCC 1702
Db 1925 CAACCGAGCCAGGCTTCCAGAGACAGAGGAGGAGGAGGAGGAGGAGGAGAGCC 1984
OY 1703 CTGCAATTCCTCTACGGCAGGTTCCGGAAGTGTGAGAGAGGCTGACCTGACATGAGAG 1762
Db 1985 CTGCAATTCCTCTACGGCAGGTTCCGGAAGTGTGAGAGAGGCTGACCTGACATGAGAG 2044
OY 1763 TGGAGAGAGGAGGCGAGGCTGAGGCTTACACATGAGCCAGGCTCCCTGACACTGAAGAG 1822
Db 2045 TGGAGAGAGGAGGCGAGGCTGAGGCTTACACATGAGCCAGGCTCCCTGACACTGAAGAG 2104
OY 1823 GATCCACAACTCTTGAAGAAACACCTGACGTGTGTGCGCACATTTCTCTCAGCT 1882
Db 2105 GATCCACAACTCTTGAAGAAACACCTGACGTGTGTGCGCACATTTCTCTCAGCT 2164
OY 1883 CGGCCCATACCGCTACCTAGAGCTGACGCTCCACCGCCCTGCACTAGGGCTCACCTCC 1942
Db 2165 CGGCCCATACCGCTACCTAGAGCTGACGCTCCACCGCCCTGCACTAGGGCTCACCTCC 2224
OY 1943 CACCCCTGTCACTACAGGCTCAACCTCTACAGCCTTAACTCCAGAGCCATGTGCTG 2002
Db 2225 CACCCCTGTCACTACAGGCTCAACCTCTACAGCCTTAACTCCAGAGCCATGTGCTG 2284
OY 2003 TCCAGGSGCTCAAGACTTTCCTAAGTGGAGTGTGTAAGGAGCTGAAGTACCTTTGGGG 2062
Db 2285 TCCAGGSGCTCAAGACTTTCCTAAGTGGAGTGTGTAAGGAGCTGAAGTACCTTTGGGG 2344
OY 2063 GCACAGAGCCCTTAACTTCTCACTTCACTTGAAGCCCTGACACATCTACCTGTGGCAGGAA 2122
Db 2345 GCACAGAGCCCTTAACTTCTCACTTCACTTGAAGCCCTGACACATCTACCTGTGGCAGGAA 2404
OY 2123 TGAAGACAGAGCTTCCGCTGCAAAAGGCTACAGGCTCCACCCCGGCGCTCCCTGCA 2182
Db 2405 TGAAGACAGAGCTTCCGCTGCAAAAGGCTACAGGCTCCACCCCGGCGCTCCCTGCA 2464
OY 2183 CCTCTGTCTCTCTCCAGTTCATCTCTGGAACACAGGCGCAGGCAACAGTGGCCCCC 2242
Db 2465 CCTCTGTCTCTCTCCAGTTCATCTCTGGAACACAGGCGCAGGCAACAGTGGCCCCC 2524
|||||

QY 2243 AAAGCAGGAGATCTCAGGCCAGCGGGAGGCTGGAAGGCTGCGAGATCGCT 2302
 Db 2525 AAAGCAGGAGATCTCAGGCCAGCGGGAGGCTGGAAGGCTGCGAGATCGCT 2584
 QY 2303 TCCCTATCCACTCCACCGGCTCCAGGCTCTTGGCTCTGCCACACCTCTGTACAC 2362
 Db 2585 TCCCTATCCACTCCACCGGCTCCAGGCTCTTGGCTCTGCCACACCTCTGTACAC 2644
 QY 2363 CACCCAGATCACAGGAGCAGCGCCAGAGATAGTCTTTTGTCTTGTGCGCTCT 2422
 Db 2645 CACCCAGATCACAGGAGCAGCGCCAGAGATAGTCTTTTGTCTTGTGCGCTCT 2704
 QY 2423 GGCTATCTAGTTTTCATATAGCTTACAGTATCTGCTTGTACTAGAAATPAAACACAT 2482
 Db 2705 GGCTATCTAGTTTTCATATAGCTTACAGTATCTGCTTGTACTAGAAATPAAACACAT 2764
 QY 2483 TTTTCATA 2489
 Db 2765 TTTTCATA 2771

RESULT 7

AAD36061
 ID AAD36061 standard; cDNA; 2718 BP.

AC AAD36061;

DT 09-AUG-2002 (first entry)

DE Human dual-specificity phosphatase 15 (DSP-15) cDNA.

XX Human; dual-specificity phosphatase 15; DSP-15; anti-allergic; cytosolic;
 KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
 KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
 KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
 KW metabolic disease; allergy; screening; chromosome 11q; gene; ss.

OS Homo sapiens.

XX Key location/Qualifiers
 FH 35.2014
 FT CDS /tag= a
 FT /product= "Human DSP-15 protein"

PN WO200224740-A2.

PD 28-MAR-2002.

PE 19-SEP-2001; 2001WO-US29406.

PR 19-SEP-2000; 2000US-233833P.

PR 18-SEP-2001; 2001US-0955732.

PA (CEPT-) CEPTYR INC.

PI Luche RM, Wei B;

DR WPI; 2002-394127/42.

DR P-PSDB; AAE22729.

XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
 PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
 PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
 PT diseases
 XX Claim 7; Fig 1; 91pp; English.

CC The invention relates to a new isolated dual-specificity phosphatase 15
 CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
 CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
 CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
 CC residues. DSP-15 polypeptides may be used to identify agents that

CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
 CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
 CC polypeptides, modulating agents, and/or polynucleotides encoding the
 CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
 CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
 CC versus host disease, autoimmune diseases, allergies, metabolic diseases,
 CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
 CC assays for modulators of enzyme activity and/or substrate binding. The
 CC present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
 CC chromosome 11q.
 SQ Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other:
 Query Match 82.0%; Score 2083.6; DB 24; Length 2718;
 Best Local Similarity 89.2%; Pred. No. 0;
 Matches 2427; Conservative 0; Mismatches 4; Indels 291; Gaps 5;
 QY 54 CCGGTGCGACGCCAGGTGCTGCGGCTGCTCCATGAGCTGTGACAGTGAAGCGGTC 113
 Db 1 CCGGTGCGACGCCAGGTGCTGCGGCTGCTCCATGAGCTGTGACAGTGAAGCGGTC 60
 QY 114 GCCCGGCGAGGCGGCTCCAGCGCCCTGGGCGCTGGACACAGGCGCTCCAGCGAAG 173
 Db 61 GCCCGGCGAGGCGGCTCCAGCGCCCTGGGCGCTGGACACAGGCGCTCCAGCGAAG 120
 QY 174 GAGTCGACTCCAGCAGAGCAGAGCTTTGCGGTGCTGCGTGGGGGCTGCTCGGACTGCA 233
 Db 121 GAGTCGACTCCAGCAGAGCAGAGCTTTGCGGTGCTGCGTGGGGGCTGCTCGGACTGCA 180
 QY 234 GGATGGAGGGGACATGATGATGACAGCAGAGCCGCTTGTGACCCACAGAGAAAGGCC 293
 Db 181 GGATGGAGGGGACATGATGATGACAGCAGAGCCGCTTGTGACCCACAGAGAAAGGCC 240
 QY 294 GAGTGAAGAGGAGCTCCAGCGGGGACACAGACTTCCGGCAGAGATCCGAGATCCCA 353
 Db 241 GAGTGAAGAGGAGCTCCAGCGGGGACACAGACTTCCGGCAGAGATCCGAGATCCCA 300
 QY 354 GAAGCAGAGGAGCAGAGCAGACCTCCTACATGATGATGAGTGAAGCGCGCAGCA 413
 Db 301 GAAGCAGAGGAGCAGAGCAGACCTCCTACATGATGATGAGTGAAGCGCGCAGCA 360
 QY 414 TGACATCGCGCTGGAGCGCCAGCTGGAAGGACCCCGGCTCCCGGCTCGCTACCTGCT 473
 Db 361 TGACATCGCGCTGGAGCGCCAGCTGGAAGGACCCCGGCTCCCGGCTCGCTACCTGCT 420
 QY 474 GGTAGTTCTACAGAGAGAGAGAGGTCTGAGCCAGATGAGACGGTCTCTCGGCGCT 533
 Db 421 GGTAGTTCTACAGAGAGAGAGAGGTCTGAGCCAGATGAGACGGTCTCTCGGCGCT 480
 QY 534 GGATTTCCCTGACAGAGCTCCCGCAGCTGACCTGAGGCGCTGCTGCTGAGAG 593
 Db 481 GGATTTCCCTGACAGAGCTCCCGCAGCTGACCTGAGGCGCTGCTGCTGAGAG 540
 QY 594 TGACACCCAGGTGTTAGATGAGAGACGGGGCTTCAGGCTGAGCTGAGTGGGCAAG 653
 Db 541 TGACACCCAGGTGTTAGATGAGAGACGGGGCTTCAGGCTGAGCTGAGTGGGCAAG 600
 QY 654 CCGATCTTCAGGCCATCTCCATCCAGACATG----- 688
 Db 601 CCGATCTTCAGGCCATCTCCATCCAGACATGTTGGCCACATCTCAGATATTGCACCA 660
 QY 689 ----- 688
 Db 661 AGCATGTAGAGCAGCTTAGAGCAGCGGCTTGTACCGGCTGGCAGTGCCTTACTGGGC 720
 QY 689 ----- 688
 Db 721 CACGCCACTACAGAGAGACTGACTCCAGACAGACTGCTCAATGATGAGAGCGGTAT 780
 QY 689 -----GTTCTCAGAACAGCA 704
 Db 781 GGCCGACTGAGGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGCTCTCTTCTGAGAACAGCA 840

Oy 705 GCAGATGAGCAGGCGATCCGTCGAGCTGAGGAAAGTGTGATGTCAGTGCAGCTGGA 764
|||
Db 841 GCAGATGAGCAGGCGATCCGTCGAGCTGAGGAAAGTGTGATGTCAGTGCAGCTGGA 900
Oy 765 GAGTGTCACTTCCAAAAGATCCGCCAGAGCTCTGAGAGCTGCGCCCTGAGGCTCCCTCCCA 824
|||
Db 901 GAGTGTCACTTCCAAAAGATCCGCCAGAGCTCTGAGAGCTGCGCCCTGAGGCTCCCTCCCA 960
Oy 825 GCAGTACCGTGACTTCAATCGACACCAATGCTGCTGGTGGCGACAGCGGAGCCGAGC 884
|||
Db 961 GCAGTACCGTGACTTCAATCGACACCAATGCTGCTGGTGGCGACAGCGGAGCCGAGC 1020
Oy 885 CTCCCGCATCTTCCCGCAGCTTACCTGGGGCTCAGAGTGGAAAGCAGCAAACTCTGAGGA 944
|||
Db 1021 CTCCCGCATCTTCCCGCAGCTTACCTGGGGCTCAGAGTGGAAAGCAGCAAACTCTGAGGA 1080
Oy 945 GCTGCAGAGGAACAGAGGTACCCACATCTTGAACATGGCCCGGAGATGTGACAACTTCTA 1004
|||
Db 1081 GCTGCAGAGGAACAGAGGTACCCACATCTTGAACATGGCCCGGAGATGTGACAACTTCTA 1140
Oy 1005 CCTTAGCGCTTCACTTACCAATGTGCGCTCTGGATGAGAGTGGCCAGCTGCT 1064
|||
Db 1141 CCTTAGCGCTTCACTTACCAATGTGCGCTCTGGATGAGAGTGGCCAGCTGCT 1200
Oy 1065 GCGGCACTGGAAGGAGAGCAGCCGCTTCAATGAGGCTGCAAGAGCAGAGGAGCCAGAGT 1124
|||
Db 1201 GCGGCACTGGAAGGAGAGCAGCCGCTTCAATGAGGCTGCAAGAGCAGAGGAGCCAGAGT 1260
Oy 1125 GCTGTTCACATGCAAGATGGGCGTCAGCCGCTCAGCGGCAAGTGTGCGCTATGCCAT 1184
|||
Db 1261 GCTGTTCACATGCAAGATGGGCGTCAGCCGCTCAGCGGCAAGTGTGCGCTATGCCAT 1320
Oy 1185 GAAGAGTACCAATGCACTTGGAGAGAGCCCTGGCCAGCTGCGAGAGCTTCCGCCAT 1244
|||
Db 1321 GAAGAGTACCAATGCACTTGGAGAGAGCCCTGGCCAGCTGCGAGAGCTTCCGCCAT 1380
Oy 1245 CGCCGCGCCCAACCTTGCTTCTGCGCAGCTGAGATCTACAGGCGATCTCTGAGC-- 1302
|||
Db 1381 CGCCGCGCCCAACCTTGCTTCTGCGCAGCTGAGATCTACAGGCGATCTCTGAGC 1440
Oy 1303 ----- 1302
Db 1441 CAGCGCCAGAGCCATGTCGTGGAGCAAAAGTGGGGTCTCCCGAGAGAGCAGCC 1500
Oy 1303 -----GCCAAGCTGAGGGTGGGGGA 1326
|||
Db 1501 AGCCCTGAAGTCTCTACACCATTCACACTCTTCGCGCAGAACTGAGGGTGGGGGA 1560
Oy 1327 GGAGAAGTGTAGGCAATGAAAGAGAGCAGGAGCCCGAAGAGAGCCGAGGGCCAGC 1386
|||
Db 1561 GGAGAAGTGTAGGCAATGAAAGAGAGCAGGAGCCCGAAGAGAGCCCTGGG--CCA 1618
Oy 1387 GGGGACGTATTAACCTCCGAGGGGTCTAGAGTCCATCACTCTTCTGAGAGCCCTCTTG 1446
|||
Db 1619 GGGGACGTATTAACCTCCGAGGGGTCTAGAGTCCATCACTCTTCTGAGAGCCCTCTT- 1677
Oy 1447 GGAGCTGAGAGCACTGATGAGACATGACATGCCAGAGGTCTTCTCTCCACGAGT 1506
|||
Db 1678 GGAGCTGAGAGCACTGATGAGACATGACATGCCAGAGGTCTTCTCTCCACGAGT 1736
Oy 1507 CTTCACATGAAGAGCCTCTGAGCCCTTCCACAGCTTGAAGAGCAGCAAGGAGGAGCAGC 1566
|||
Db 1737 CTTCACATGAAGAGCCTCTGAGCCCTTCCACAGCTTGAAGAGCAGCAAGGAGGAGCAGC 1796
Oy 1567 AGGTGACAGGGGGCTCAGCTGCGCTGAAGTCCCGCAGTCACTGCTTACCTTCAGG 1626
|||
Db 1797 AGGTGACAGGGGGCTCAGCTGCGCTGAAGTCCCGCAGTCACTGCTTACCTTCAGG 1856
Oy 1627 GCAGTGCCTGTGGCCCAACCGGAGAGGCTTCCAGAGCAGAGAGAGGAGGAGGAGG 1686
|||
Db 1857 GCAGTGCCTGTGGCCCAACCGGAGAGGCTTCCAGAGCAGAGAGAGGAGGAGGAGG 1916

Oy 1687 AGGGCAGGAGAGCCCTGCATTTCTCTACGCCAGGTTCGGAAGGTGGAGACAGG 1746
|||
Db 1917 AGGGCAGGAGAGCCCTGCATTTCTCTACGCCAGGTTCGGAAGGTGGAGACAGG 1976
Oy 1747 CCAGGTCATGATGAGTGGAGAGAGGGGAGGCTTGAGCCCTCAGACATGCCAGCTC 1806
|||
Db 1977 CCAGGTCATGATGAGTGGAGAGAGGGGAGGCTTGAGCCCTCAGACATGCCAGCTC 2036
Oy 1807 CCTTACACTGAAGAGATCCCAACATCTCTTGGAGAAACACCCCTACGTCTGTCCGCA 1866
|||
Db 2037 CCTTACACTGAAGAGATCCCAACATCTCTTGGAGAAACACCCCTACGTCTGTCCGCA 2096
Oy 1867 CACATTTCTTCAGCTCCGCCCATATCCGCTACAGAGCCCTCAGCCACCCCTGCA 1926
|||
Db 2097 CACATTTCTTCAGCTCCGCCCATATCCGCTACAGAGCCCTCAGCCCTGCA 2156
Oy 1927 CTAGGCGCTCACTCCACCCCTGTCACTACAGCTTCACTCTACAGCCCTTAAGTCCA 1986
|||
Db 2157 CTAGGCGCTCACTCCACCCCTGTCACTACAGCTTCACTCTACAGCCCTTAAGTCCA 2216
Oy 1987 GGCCCATGTCTGCTGTCCAGAGGCTCAAGACTTCTTACCTGAGATGTGTAGAGGACT 2046
|||
Db 2217 GGCCCATGTCTGCTGTCCAGAGGCTCAAGACTTCTTACCTGAGATGTGTAGAGGACT 2276
Oy 2047 GAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCACTTCACTTCACTTCACTTCACT 2106
|||
Db 2277 GAAGTACCTTTGGGGGCAACAGCACCCCTAGTTTCACTTCACTTCACTTCACTTCACT 2336
Oy 2107 CACCTGTGGCAGGAATGAAAAAGAGCTTCCCGTCAAAAAAGGCTCAGCCCTCCACCC 2166
|||
Db 2337 CACCTGTGGCAGGAATGAAAAAGAGCTTCCCGTCAAAAAAGGCTCAGCCCTCCACCC 2396
Oy 2167 CCGCCCTCTCTGTGACCTCTCTGTCTCTCTCCAGTTCACTTCTGGAACAGCAGGCGCAG 2226
|||
Db 2397 CCGCCCTCTCTGTGACCTCTCTGTCTCTCTCCAGTTCACTTCTGGAACAGCAGGCGCAG 2456
Oy 2227 GCAACAGATGGGCCCAAGAGCAGAGCAGAGATCTCAGGCGCCAGCGGGAGGCTGGA 2286
|||
Db 2457 GCAACAGATGGGCCCAAGAGCAGAGCAGAGATCTCAGGCGCCAGCGGGAGGCTGGA 2516
Oy 2287 GGGCTGCGAGATCGCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 2346
|||
Db 2517 GGGCTGCGAGATCGCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 2576
Oy 2347 AGACCTCTGTGACACAGCAGCAGATCAGAGGAGCAGCAGGCGCAGAGATGCTCTTTT 2406
|||
Db 2577 AGACCTCTGTGACACAGCAGCAGATCAGAGGAGCAGCAGGCGCAGAGATGCTCTTTT 2636
Oy 2407 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2466
|||
Db 2637 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2696
Oy 2467 GAGAAATTAACACATTTTCAT 2488
|||
Db 2697 GAGAAATTAACACATTTTCAT 2718
|||
RESULT 8
AAH14722
ID AAH14722 standard; cDNA; 1755 BP.
AC AAH14722;
AC xx
AC xx
DT 26-JUN-2001 (first entry)
DE Human cDNA sequence SEQ ID NO:12452.
DE xx
DE xx
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW xx
OS Homo sapiens.
OS xx
PN EPI074617-A2.
PN xx

PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PT
 PS Claim 8; SEQ ID 12452; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH3633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH3683 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;

Query Match 51.3%; Score 1303.6; DB 22; Length 1755;
 Best Local Similarity 94.2%; Pred. No. 8e-238;
 Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;

QY 979 ATGGGCGGGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCCCTC 1038
 DB 322 ATGGCGCCGGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGGCCCTC 381
 QY 1039 TGGATGAGAGATCGGCCAGCTGCTGCGCACTGGAAGGAGAGACCGCTTCAATTAG 1098
 DB 382 TGGATGAGAGATCGGCCAGCTGCTGCGCACTGGAAGGAGAGACCGCTTCAATTAG 441
 QY 1099 GCTGCAAGAGACAGAGGAGACCGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
 DB 442 GCTGCAAGAGACAGAGGAGACCGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
 QY 1159 GCGGCGAAG 1218
 DB 502 GCGGCGAAG 561
 QY 1219 GCGGCGAAG 1278
 DB 562 GCGGCGAAG 621

QY 1279 CAGATCTACAGGCGCATCTGAGCGCCAGAACCTGAGGGGTGTGGGAGAGAGAGTTGT 1338
 DB 622 CAGATCTACAGGCGCATCTGAGCGCCAGAACCTGAGGGGTGTGGGAGAGAGAGTTGT 681
 QY 1339 AGGATGAG 1398
 DB 682 AGGATGAG 739
 QY 1399 AACCTCGAAGGAGTCAAGAGTCAATCACTCTTCTGAGAGCCCTCTTGGAGCTGAGAG 1458
 DB 740 AACCTCGAAGGAGTCAAGAGTCAATCACTCTTCTGAGAGCCCTCTTGGAGCTGAGAG 798
 QY 1459 CACCTAGTAGACAGATGACATGACAGAGCTTCTCTTCCACAGAGTTCATCATGAG 1518
 DB 799 CACCTAG-AGACAGATGACATGACAGAGCTTCTCTTCCACAGAGTTCATCATGAG 857
 QY 1519 AGCCTGACAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1578
 DB 858 AGCCTGACAGCCCTTCCACAGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 917
 QY 1579 GGCCTCAGCCTGCGCTGAAGTCCCGCAGTCACTGAGTTACCTCCAGGAGAGAGAG 1638
 DB 918 GGCCTCAGCCTGCGCTGAAGTCCCGCAGTCACTGAGTTACCTCCAGGAGAGAGAG 977
 QY 1639 TGGCCACCGGAG 1698
 DB 978 TGGCCACCGGAG 1037
 QY 1699 AGCCTGATTTCTCTAAGCCAGGTTCCGGAAGGTGTGAGAGAGAGAGAGAGAGAG 1758
 DB 1038 AGCCTGATTTCTCTAAGCCAGGTTCCGGAAGGTGTGAGAGAGAGAGAGAGAGAG 1097
 QY 1759 ACAGTGGAG 1818
 DB 1098 ACAGTGGAG 1157
 QY 1819 AGAGATCACACACCTCTTGGAGAACACCTCAGCTGTGTTGGCGACACATCTCTCTC 1878
 DB 1158 AGAGATCACACACCTCTTGGAGAACACCTCAGCTGTGTTGGCGACACATCTCTCTC 1217
 QY 1879 AGCTCGGCGCATACCGGTCTACATAGAGCTTACCTCCAGCCCTGTCACTAGAGAG 1938
 DB 1218 AGCTCGGCGCATACCGGTCTACATAGAGCTTACCTCCAGCCCTGTCACTAGAGAG 1277
 QY 1939 CTCCAGCCCTGTCTACATAGAGCTTACCTCCAGCCCTGTCACTAGAGAGAGAGAG 1998
 DB 1278 CTCCAGCCCTGTCTACATAGAGCTTACCTCCAGCCCTGTCACTAGAGAGAGAGAG 1337
 QY 1999 CTTGTCAGAGGCTCAACAGCTTCTTAACCTGAGATGTGTAGAGAGAGAGAGAGAG 2058
 DB 1338 CTTGTCAGAGGCTCAACAGCTTCTTAACCTGAGATGTGTAGAGAGAGAGAGAGAG 1397
 QY 2059 GGGGGAACAGACACCTAGTTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 2118
 DB 1398 GGGGGAACAGACACCTAGTTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1434
 QY 2119 GGAATGAAAACAGAGCTTCCGCTGCAAAAAGAGGTCAAGCTTCCAGCCCGCCCTTCC 2178
 DB 1435 -----CC 1437
 QY 2179 TGCACCT 2238
 DB 1438 TGCACCT 1497
 QY 2239 CCCCAGAGAGAGAGAGATCTCTAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2298
 DB 1498 CCCCAGAGAGAGAGAGATCTCTAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
 QY 2299 CGCTTCCCTATCTACCTCTCAAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2358
 DB 1558 CGCTTCCCTATCTACCTCTCAAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1617
 QY 2359 ACACAG 2418

|||||
Db 1618 ACACACGCGACATGATCAGAGGCGACGACAGATAGCTCTTTGTCTCTTCTGCGC 1677
OY 2419 CTTCTGGCTAGTACGTTTTCATACGCTTACATATCTGCTTTGACGAGAAATAAAC 2478
Db 1678 CTTCTGGCTAGTACGTTTTCATACGCTTACATATCTGCTTTGACGAGAAATAAAC 1737
OY 2479 ACATTTTCATATTGGTT 2496
Db 1738 ACATTTTCATATTGGTT 1755

RESULT 9
ABL40801
ID ABL40801 standard; DNA; 1755 BP.
XX
AC ABL40801;
XX
03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KM Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KM antiasthmatic; antiidiabetic; anorectic; cytoplastic; cardiant; human;
KM antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KM neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KM antiallergic; dermatological; vulnery; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2.
XX
PD 14-MAR-2002.
XX
PE 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX
PI Liou J;
XX
DR WPI; 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases -
XX
PS Disclosure; Fig 4; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
XX

Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;
SQ
Query Match 51.3%; Score 1303.6; DB 24; Length 1755;
Best Local Similarity 94.2%; Pred. No. 8e-238;
Matches 1430; Conservative 0; Mismatches 4; Indels 84; Gaps 4;

OY 979 ATGGCCGGGAGATTGACAACTTCTACCTTACGACCTTACACATATGCGCCTC 1038
Db 322 ATGGCCGGGAGATTGACAACTTCTACCTTACGACCTTACACATATGCGCCTC 381
OY 1039 TGGGATGAGAGTCCGCCAGCTGCTGCCGCACTGGAAGAGACGCCCTTCAATTAG 1098
Db 382 TGGGATGAGAGTCCGCCAGCTGCTGCCGCACTGGAAGAGACGCCCTTCAATTAG 441
OY 1099 GCTGCAAGAGCAGAGGCAACCCAGCTGCTGCTCCACTGCAAGATGGGCTTACGCCCTCA 1158
Db 442 GCTGCAAGAGCAGAGGCAACCCAGCTGCTGCTCCACTGCAAGATGGGCTTACGCCCTCA 501
OY 1159 GCGGCCACAGTGTGGGCTATGCTGCAAGCAGTACGAATGAGCGCTGAGAGGCCCTG 1218
Db 502 GCGGCCACAGTGTGGGCTATGCTGCAAGCAGTACGAATGAGCGCTGAGAGGCCCTG 561
OY 1219 CGCCACGTGCAAGAGCTCCGCCATCGCCGCCAACCCTGAGCTTCTGCGCCAGCTG 1278
Db 562 CGCCACGTGCAAGAGCTCCGCCATCGCCGCCAACCCTGAGCTTCTGCGCCAGCTG 621
OY 1279 CAGATCTACCCAGGAGATCTCTGAGGCGCAAGAACCTTAGGGTGGTGGGAGGAAGTGT 1338
Db 622 CAGATCTACCCAGGAGATCTCTGAGGCGCAAGAACCTTAGGGTGGTGGGAGGAAGTGT 681
OY 1339 AGGCATGGAAGAGAGCCAGAGCCAGCCCGAAGAGAGCTGGGGGAGGAGCTATA 1398
Db 682 AGGCATGGAAGAGAGCCAGAGCCAGCCCGAAGAGAGCTGGGGGAGGAGCTATA 739
OY 1399 AACCTTCGAGGGGTATGAGTGCATCAGTCTTGTGAGACCCTCTTGGGAGCTGAGAG 1458
Db 740 AACCTTCGAGGGGTATGAGTGCATCAGTCTTGTGAGACCCTCTTGGGAGCTGAGAG 798
OY 1459 CAGCTCAGTAGAGCAGTACATGACATGACAGAGTCTTCTCCAGAGTCTTACATGAG 1518
Db 799 CAGCTCAGTAGAGCAGTACATGACATGACAGAGTCTTCTCCAGAGTCTTACATGAG 857
OY 1519 AGCCTTCGAGAGCCCTTCCACAGTTCGAGAGAGCCAGAGAGGAGGAGGAGGAG 1578
Db 858 AGCCTTCGAGAGCCCTTCCACAGTTCGAGAGAGCCAGAGAGGAGGAGGAGGAGGAG 917
OY 1579 GGCTTCAGGCTGCTCTGAGAGTCCGCCAGTCAAGTGTACCTTCCAGAGGAGCTGCTG 1638
Db 918 GGCTTCAGGCTGCTCTGAGAGTCCGCCAGTCAAGTGTACCTTCCAGAGGAGCTGCTG 977
OY 1639 TGGCCACCCGAGCCAGGCTTCCAGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAG 1698
Db 978 TGGCCACCCGAGCCAGGCTTCCAGAGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAG 1037
OY 1699 AGCCTTCATTTCTCTAGAGCTCCGAGAGTTCGGAAGGTGAGAGAGCAGCAGCTGATG 1758
Db 1038 AGCCTTCATTTCTCTAGAGCTCCGAGAGTTCGGAAGGTGAGAGAGCAGCAGCTGATG 1097
OY 1759 ACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1818
Db 1098 ACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1157
OY 1819 AGAGATCCACACACTCTTTGGAGAAACACCTCAGCTGTGTTGGCGACACATTTCTCTC 1878
Db 1158 AGAGATCCACACACTCTTTGGAGAAACACCTCAGCTGTGTTGGCGACACATTTCTCTC 1217
OY 1879 AGCTCGGCCCCATACCCGATACAGCTGACCTGACCCGCTTCACTACAGGCTTAC 1938
Db 1218 AGCTCGGCCCCATACCCGATACAGCTGACCTGACCCGCTTCACTACAGGCTTAC 1277
OY 1939 CTCCACACCCCTGCTACTACAGGCTTCACTACAGGCTTAACTTCAAGTCCAGGCCATGCTG 1998
Db 1278 CTCCACACCCCTGCTACTACAGGCTTCACTACAGGCTTAACTTCAAGTCCAGGCCATGCTG 1337

Oy	1999	CCTGTCCAGGGCTCAAGACTTTCCTTAAC TGATGCGTAGAGGGAGCTGAAGGTAACCTTT	2058
Db	1338	CCTGTCCAGGGCTCAAGACTTTCCTTAAC TGATGCGTAGAGGGAGCTGAAGGTAACCTTT	1397
Oy	2059	GGGGGGCAACAGCACACCCTAGTTTCATTCTGCAC TTAGCCCCCTCACA CTACACCTGTGGCAC	2118
Db	1398	GGGGGGCAACAGCACACCCTAGTTTCATTCTGCAC TTAGCCCCCTCACA CTACACCTGTGGCAC	1434
Oy	2119	GGATGAAACAGAGCTTCCCCTGCAAAAAGSTGAC GGCCTCCACCCCGCCCTGCC	2178
Db	1435	----- ---CC 1437	
Oy	2179	TGCACCTCTGTCTCTCCAGTTTCATCTCTGGAAC CAGCCAGGCAGCAACAGTGCC	2238
Db	1438	TGCACCTCTGTCTCTCCAGTTTCATCTCTGGAAC CAGCCAGGCAGCAACAGTGCC	1497
Oy	2239	CCCCAAGAAGCAGAGATCTCTAGGCCCCACG CCGGGGAGGCTGGAAGGCTGGCAGAT	2298
Db	1498	CCCCAAGAAGCAGAGATCTCTAGGCCCCACG CCGGGGAGGCTGGAAGGCTGGCAGAT	1557
Oy	2299	CGCTTCCCTCATCCACCTCCACCGGTCAGAG GTTTTGCTGTCTGCCAGACCTCTGTG	2358
Db	1558	CGCTTCCCTCATCCACCTCCACCGGTCAGAG GTTTTGCTGTCTGCCAGACCTCTGTG	1617
Oy	2359	ACACCCAGCCAGATCACAGGGCACACAGGCC AGAGATAGTCTTTTGTCTCTTGTGCC	2418
Db	1618	ACACCCAGCCAGATCACAGGGCACACAGGCC AGAGATAGTCTTTTGTCTCTTGTGCC	1677
Oy	2419	CTGTGGCTAGTCAGTTTTTTCATAGCCCTTTC AAGATCTGGCTTGTACTGAGAAATAAAC	2478
Db	1678	CTGTGGCTAGTCAGTTTTTTCATAGCCCTTTC AAGATCTGGCTTGTACTGAGAAATAAAC	1737
Oy	2479	ACATTTTCATATTGGTT 2496 	
Db	1738	ACATTTTCATATTGGTT 1755	
<hr/>			
RESULT 10			
ABL40805			
ID	ABL40805	standard; cDNA; 2322 BP.	
XX	AC		
XX	ABL40805;		
XX	03-JUL-2002	(first entry)	
De	Human	MAP kinase phosphatase-like enzyme encoding cDNA.	
XX	Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;		
KW	antidiabetic; anorectic; cytostatic; cardiant; human;		
KW	antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;		
KW	neuroleptic; anticonvulsant; anti-HIV; antirhythmic; hypotensive;		
KW	antiallergic; dermatological; vulnerary; gene therapy; ss.		
XX	Homo sapiens.		
XS			
Key	Location/Qualifiers		
FH	1..2322		
FT	/+tag= a		
FT	/product= "MAP kinase phosphatase-like enzyme"		
FT	/note= "contains internal codon deletions"		
FT	/transl_except= "(pos: 156..157, aa: Leu)"		
FT	/note= "there is an apparent one codon deletion which		
FT	alters the reading frame"		
FT	/transl_except= "(pos: 180..181, aa: Arg)"		
FT	/note= "there is an apparent one codon deletion which		
FT	alters the reading frame"		
FT	/transl_except= "(pos: 414..415, aa: Val)"		
FT	/note= "there is an apparent one codon deletion which		
FT	alters the reading frame"		
FT	/transl_except= "(pos: 1197..1198, aa: Tyr)"		
FT	/note= "there is an apparent one codon deletion which		
FT	alters the reading frame"		

[illegible]

420 CCCTCTGCA--GCCAGCTGAGGACCCGGCTCCCGGCTCCGCTACCTGC----- 472
443 CCGACTTGGAGCCCAAGCTGAGGACCCCGGGCTCCCGGGATCCGGATACCTTGT 502
473 TGTGATTTCTACACGAGAAG-GAGAGGTCTGAG-CCAGATGAGAGGCTCTCTGG 530
503 TGTGATTTCTACACGAGAAGGAGAGGATGAGCCAGATGAGAGGCTCTCTGG 562
531 -CGTGATTTCTACACGAGCTCCCGAGCTGCACCTGGGCTGGCTTGGCCCTCT 589
563 ACGTGATTTCTCTACAGCACTCTCCAGCTGCACCTGGGCTGGCTTGGCCCTCT 622
590 GAGTGACACCCAGGCTTACTTACTAGTGAAGAGGAGGCTTGAAGCTGAGCTGAGG 649
623 GGAGTGACACCCAGGCTTACTTACTAGTGAAGAGGAGGCTTGAAGCTGAGCTGAGG 682
650 AAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGT----- 688
683 AAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGTATTGC 742
689 ----- 688
743 ACCAAGCATGTGAGGCACTCTAGGCAAGGCTTGTACCGGGTGGCAGTCCCTACCT 802
689 ----- 688
803 GGGCAGCCACTACAGGAGAGACTGAACTCCGAACAGAGCTGCCTCAATGATGAGG 862
689 ----- 700
863 CTATGCGCAGCTGAGTCTCTGCGGCTCCAGCGCCAGCCCTGGCGTCCAGAAC 922
701 AGGACAGATGAGAGGAGCGGATCCGCTGAGCTGGAAGTGTGATGATGATGAC 760
923 AGGACAGATGAGAGGAGCGGATCCGCTGAGCTGGAAGTGTGATGATGATGAC 982
761 TGGAGAGTGTCTTCCAAAGAGATCCGCGAGGCTGTGAGCTGGCGCTGGCCCTCC 820
983 TGGAGAGTGTCTTCCAAAGAGATCCGCGAGGCTGTGAGCTGGCGCTGGCCCTCC 1042
821 TCCAGCAGTACCGTGTCTTCCAAAGAGATCCGCGAGGCTGTGAGCTGGCGCTGG 880
1043 TCCAGCAGTACCGTGTCTTCCAAAGAGATCCGCGAGGCTGTGAGCTGGCGCTGG 1102
881 GAGCTCCCGATCTTCCCGACCTCTACCTGAGCTGAGTGAAGCCAGAACTGG 940
1103 GAGCTCCCGATCTTCCCGACCTCTACCTGAGCTGAGTGAAGCCAGAACTGG 1162
941 AGGAGCTGAGAGAAAGGCTGAGCCACATCTTGAACATGGCCCGGAGATGACACT 1000
1163 AGGAGCTGAGAGAAAGGCTGAGCCACATCTTGAACATGGCCCGGAGATGACACT 1219
1001 TCTACCTGAGCGCTTCACTACCAATGTGCGCTCTGGAGTGAAGAGTGGCCAGC 1060
1220 TCTACCTGAGCGCTTCACTACCAATGTGCGCTCTGGAGTGAAGAGTGGCCAGC 1279
1061 TGTGCGCGCATGGAAG 1120
1280 TGTGCGCGCATGGAAG 1339
1121 ACGTGCTGGTCACTGCAAGATGGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAG 1180
1340 ACGTGCTGGTCACTGCAAGATGGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAG 1399
1181 CCATGAGCAGTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1240
1400 CCATGAGCAGTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1459
1241 CCATGAGCAGTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1300
1460 CCATGAGCAGTACGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1519

1301 CG----- 1302
1520 GGGCAGCGCCAGAGCCATGTCTGGAGCAGAGAAGTGGTGGGTCTCCCAAGAGAGC 1579
1303 -----GCCAGACTGAGGCTG 1322
1580 ACCAGCCCTGAAGTCTCTACACCATTCACCTCTTCCGCGCAGAACCTGAGGGGTG 1639
1323 GGGAGAGAGGTTGTAGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1382
1640 GGGAGAGAGGTTGTAGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1698
1383 CCACGGGACGCTATTAACCTCCAGGGGCTCATGAGTTCATAGTCTTGGAGCCCTC 1442
1699 -CCAGGCGCAGCTATTAACCTCCAGGGGCTCATGAGTTCATAGTCTTGGAGCCCTC 1757
1443 CTTGGAGCTGAGAGACACCTAGTAGACCACTGACATGCCAGAGAGTCTTCTCCAC 1502
1758 CTT-GGAGCTGAGAGACACCTAG-AGACAGTGAACATGCCAGAGTCTTCTCCAC 1815
1503 GAGTCTTACATGAAGAGAGCTCTGAGCCCTTCCACAGCTTGAAGAGAGAGAGAGAG 1562
1816 GAGTCTTACATGAAGAGAGCTCTGAGCCCTTCCACAGCTTGAAGAGAGAGAGAGAG 1875
1563 CAGCAGTGTGAGAGAGGAGGCTCAGCTGCTGAGTCCCGCAGTCACTAGTGGTACCTC 1622
1876 CAGCAGTGTGAGAGAGGAGGCTCAGCTGCTGAGTCCCGCAGTCACTAGTGGTACCTC 1935
1623 CAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682
1936 CAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1995
1683 GGGCAGGGGAG 1742
1996 GGGCAGGGGAG 2055
1743 CAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802
2056 CAGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2113
1803 GCTCCCTGACACTGAGAGAGATCCAACTCTTGGAGAAACACCTTCAGCTGTGTC 1862
2114 GCTCCCTGACACTGAGAGAGATCCAACTCTTGGAGAAACACCTTCAGCTGTGTC 2173
1863 GGCACACATCTCTCAGTCCGCGCCCATACCTGCTCAGTCACTAGCTCACTCCACCT 1922
2174 GGCACACATCTCTCAGTCCGCGCCCATACCTGCTCAGTCACTAGCTCACTCCACCT 2233
1923 GTCACAGGAGCTCACTCCACCTGCTCAGTCACTAGCTCACTAGCTCACTAGCT 1982
2234 GTCACAGGAGCTCACTCCACCTGCTCAGTCACTAGCTCACTAGCTCACTAGCT 2293
1983 CCGAGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2011
2294 CCGAGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2322

RESULT 11
ABNS9832
ID ABNS9832 standard; cDNA: 2061 BP.
XX
AC ABNS9832;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 243.
XX
KW Human; antianaemic; vulnerey; antiinflammatory; immunomodulator;
KW antineoplastic; cerebroprotective; cytoskeletal; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.

XX WO200222660-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US26015.
XX 11-SEP-2000; 2000US-0659671.
XX (HYSEQ-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX MPI: 2002-292408/33.
XX P-PSDB; ABB97419.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Claim 1; SEQ ID NO 243; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
XX novel human proteins. These were isolated from expressed sequences tags
XX (ESTs). They can be used to stimulate cell growth, to regulate
XX hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX e.g. in burn treatment, to regulate the immune system e.g. to treat
XX multiple sclerosis, to regulate actinin or inhibin e.g. to treat
XX infertility, to regulate haemostasis or thrombolysis e.g. to treat
XX stroke and cancer, to screen for drugs, to treat inflammatory conditions
XX e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
XX Parkinson's disease. The present sequence is a coding sequence of the
XX invention.
XX
XX Sequence 2061 BP; 415 A; 672 C; 605 G; 369 T; 0 other:
SQ
Query Match 47.98; Score 1216.4; DB 24; Length 2061;
Best Local Similarity 74.68; Pred. No. 2.8e-221;
Matches 1875; Conservative 0; Mismatches 126; Indels 513; Gaps 7;
1 COTGATCTCGGGGCTCCAGAGACTGTCCCGGGGTTGAGGGAAGGGCCCTCCCGGTGC 60
Db 61 CTTGATCTCGGGGCTCCAGAGACTGT-CCGCGGGGTTGAGGGAAGGGCCCTCCCGGTGC 119
QY 61 CAGCCAGAGTGTCTCGGGGCTGTGCATGCGCTGTGCATGAGCCGTTGCGCCCG 120
Db 120 CAGCCAGAGTGTCTCGGGGCTGTGCATGCGCTGTGCATGAGCCGTTGCGCCCG 179
QY 121 GCGACGGGCGCTCCAGCGCCCTGGGGCCCTGGACACAGGGCGTCCAGGAAGAGTGA 180
Db 180 GCGACGGGCGCTCCAGCGCCCTGGGGCCCTGGACACAGGGCGTCCAGGAAGAGTGA 239
QY 181 CTCACGCAAGGACAGAGCTTTCGGTGCCTGTGGGGCTGTCTGGAGTGAAGATGA 240
Db 240 CTCACGCAAGGACAGAGCTTTCGGTGCCTGTGGGGCTGTCTGGAGTGAAGATGA 299
QY 241 GGGGCAATGATGATGACAGAGAGCCAGTTCTGAGCCAAACAGAGAGGCCCGCGATGAG 300
Db 300 GGGGCAATGATGATGACAGAGAGCCAGTTCTGAGCCAAACAGAGAGGCCCGCGATGAG 359
QY 301 GAGAGCTTCACGGGGACACAGACTTTCGGGCAAGAGATCCAGAGTCCCAAGAACAG 360
Db 360 GAGAGCTTCACGGGGACACAGACTTTCGGGCAAGAGATCCAGAGTCCCAAGAACAG 419
QY 361 GAGGAGCAGAGGACACCTGACCTCATGATGATGATGATGATGATGATGATGATGATG 420
Db 420 GAGGAGCAGAGGACACCTGACCTCATGATGATGATGATGATGATGATGATGATGATG 479
QY 421 GCGCTGGAGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGAT 480
Db 480 GCGCTGGAGAGCCAGCTGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGAT 539
QY 481 TCTACACGAGAGGAGAGGTCTGAGCCAGATGAGACGGTCTCTGCGGCTGATTTTC 540

Db 540 TCTACACGAGAGGAGAGGTCTGAGCCAGATGAGAGGCTCTCTGCGGCTGATTTTC 599
QY 541 CCGTACAGACAGTCCCGGAGCTGACACCTGGGACCTGGTCTTGGCCCTTGAGATGACAC 600
Db 600 CCGTACAGACAGTCCCGGAGCTGACACCTGGGACCTGGTCTTGGCCCTTGAGATGACAC 659
QY 601 CAGGTGTACTAGATGAGAGCGGGGCTTCAGCTGACGTCTGTTGGCCAAAGCCGATC 660
Db 660 CAGGTGTACTAGATGAGAGCGGGGCTTCAGCTGACGTCTGTTGGCCAAAGCCGATC 719
QY 661 TTCAAGCCCATCTCCATCCAGACCATGTGTCTGAGAAACAGAGACATGAGAGCGG 720
Db 720 TTCAAGCCCATCTCCATCCAGACCATGTGTGTGAGCACTTCAGG----- 762
QY 721 ATCCGTGTGAGCTGTGAAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 780
Db 763 ----- 762
QY 781 GAGATCCCGGAGCTGTGAGCTGCGCTGGGGCTCCCTCCAGACAGTACCGTACCTTC 840
Db 783 -----TATTGC 768
QY 841 ATGCAACAACAGATCTGCTGTGTGTCACAGCGGAGACCGAGCTCCCGCATCTTCCC 900
Db 789 ACCAAGCATGTGAGGACCTTAGGACAGCGGCTTGTACCGGG----- 811
QY 901 CACCTCTACTGTGCTCAGAGTGAACGACGACCAACTGAGAGAGTGCAGAGAACAG 960
Db 812 ----- 811
QY 961 GTACCCACATCTTAAACATGAGCCCGGAGATGATGACAACTTACCTGAGGCTTACAC 1020
Db 812 -----TGACAGTGCCTCACTACCTGAGGACGACCATGACAGAGAGACTGAGAC 857
QY 1021 TACCAATGTCGCTGCTGGAGTGAAGAGTGGCCACAGCTCTCGGACATGAGAAAGAG 1080
Db 838 TCCGACACAGAGTGTGCTCAATGA----- 880
QY 1081 ACGCACCGCTTCAATTGAGGCTCAAGAGACACAGAGGACCCACAGTGTGTCCATGCAAG 1140
Db 881 ----- 880
QY 1141 ATGGGCTGAGCGCTCAGCGGACAGTGTGCTGCTATGCTATGATGATGATGATGATG 1200
Db 881 ----- 880
QY 1201 AGCTTGAGACAGCCCTGCGGACGACGAGCTGCGGACCATGCGCCCGCAACCT 1260
Db 881 -----GTGACGCGCTATGCGGACCTGAGTCTGTGCGGCTTCCAGCGCGGAGCTTGC 935
QY 1261 GGCTTCTTCGCGGACGTGAGATCTACAGGGCATCTGACGGCCAGAACCTTGAGGTTGG 1320
Db 936 GGCTTCTTCAGACAGAGAGAGATGAGAGCGAGGATC-----GTGC 976
QY 1321 TGGGAGAGAGAGGTTGTAGGCTGAGAGAGACAGGAGCGCCCAAGAGAGGCTTGG 1380
Db 977 TGAGCTGTGAGAAAGT----- 991
QY 1381 GGCACAGGGGACAGTATAACCTCCGAGGGGTATGAGTCCATGATGATGATGATGATG 1440
Db 992 ----- 991
QY 1441 TCCCTGGAGCTGAGAGACCTCACTAGTAGACAGTACATGACAGAGTCTTCTCTCC 1500
Db 992 -----GTGAGAGCTGAGAGAGACCTCAG-AGACAGTACATGCAAGAGTCTTCTCTCC 1047
QY 1501 ACGAGTCTTACATGAGAGCTTTCAGACCTTCCACAGCTTTCAGAGACAGAGAGAG 1560
Db 1048 ACGAGTCTTACATGAGAGCTTTCAGACCTTCCACAGCTTTCAGAGAGAGAGAGAG 1107
QY 1561 GCCAGAGTGTGACAGAGGGGCTCAGGCTTGCCTGAGTGTCCGCGCAGTGTGAGTTACC 1620

Db	1108	GCACGACAGTGGACACAGGGGGCCCTCAGCTCGTCCCTGAAGTCCGCCACATGATGTTATACC	1167
Oy	1621	TCACGGCCAGTGCCTCGTGTGGCCAAACCGGACCCAGGCTTCCAGGAGCAGAGCAGGGGC	1680
Db	1168	TCACGGGCGACATGCCGTGTGGCCAAACCGGACCCAGGCTTCCAGGAGCAGAGCAGGGGC	1227
Oy	1681	AGGGGCGAGGGGCAAGGAGAACCTCGCATTTTCTCTACGCCCAAGTTTCCGGAAAGTGTGGA	1740
Db	1228	AGGGGCGAGGGGCAAGGAGAACCTCGCATTTTCTCTACGCCCAAGTTTCCGGAAAGTGTGGA	1287
Oy	1741	GACAGGCGCAGCGATGACATGACATGGAGAGAGAGGGGAGGCGCTGCAGCCCTCACACATGGCC	1800
Db	1288	GACAGGCGCAGCGATGACATGACATGGAGAGAGAGGGGAGGCGCTGCAGCCCTCACACATGGCC	1347
Oy	1801	ACGCTCCCTCTGACACTGAAGAGGAGATCACACACTCCTTGAGAAACACCCCTCAGCTGTGT	1860
Db	1348	ACGCTCCCTCTGACACTGAAGAGGAGATCACACACTCCTTGAGAAACACCCCTCAGCTGTGT	1407
Oy	1861	GCGCGCACATTCCTCTCAGCTTCGCCCCATACCCGTCACTACAGCGTCACCTCCACCC	1920
Db	1408	GCGCGCACATTCCTCTCAGCTTCGCCCCATACCCGTCACTACAGCGTCACCTCCACCC	1457
Oy	1921	CTGTGACTAGGGGCTCACCCTCCACCCCGTGTACTACACACCTCAGCTCCTACAGCTTAA	1980
Db	1468	CTGTGACTAGGGGCTCACCCTCCACCCCGTGTACTACACACCTCAGCTCCTACAGCTTAA	1527
Oy	1981	GTCGCCAGGCCCATGTGCTGTGCACAAAGGCTCAAGACTTTTAACTGGAGTGTGTGA	2040
Db	1528	GTCGCCAGGCCCATGTGCTGTGCACAAAGGCTCAAGACTTTTAACTGGAGTGTGTGA	1587
Oy	2041	GGGAGCTGAAGTACCTTTTGGGGGCAACAGCACCCCTAGTTTCAATTCTCAACTCTAGCCCTG	2100
Db	1588	GGGAGCTGAAGTACCTTTTGGGGGCAACAGCACCCCTAGTTTCAATTCTCAACTCTAGCCCTG	1647
Oy	2101	CACACTCACTGTGTGGCACGGAATGAAGAAACAGACTTCCCGTCAAAAAGGTCACGCTC	2160
Db	1648	CACACTCACTGTGTGGCACGGAATGAAGAAACAGACTTCCCGTCAAAAAGGTCACGCTC	1707
Oy	2161	CCACCCCCCGCCCTCCCTCCTCAGCTCCTGTGCCCTCCAGTTCATTCTGGACAGCGCA	2220
Db	1708	CCACCCCCCGCCCTCCCTCCTCAGCTCCTGTGCCCTCCAGTTCATTCTGGACAGCGCA	1767
Oy	2221	GGCCAGGCAACCAATGTGGCCCCCAAGAGGCAAGGAGATCTCAGGCCCAAGCCGCGGGAGG	2280
Db	1768	GGCCAGGCAACCAATGTGGCCCCCAAGAGGCAAGGAGATCTCAGGCCCAAGCCGCGGGAGG	1827
Oy	2281	CTGGAAGGGCTGGCAGATGCGTTCCTCTCATCCACTCAGCCGGTCAAGTCTTGTCTCT	2340
Db	1828	CTGGAAGGGCTGGGAGATGCGTTCCTCTCATCCACTCAGCCGGTCAAGTCTTGTCTCT	1887
Oy	2341	GTCCCCAGACCTCTGTGACATCCAGCCGATACAGGGGACACAGGCCAGAGATGTT	2400
Db	1888	GTCCCCAGACCTCTGTGACATCCAGCCGATACAGGGGACACAGGCCAGAGATGTT	1947
Oy	2401	CTTTTGTCTTTCTGTGGCCTGTGGCTAGTACAGTTTTTTCATAGCCTTACAGTATCGGCT	2460
Db	1948	CTTTTGTCTTTCTGTGGCCTGTGGCTAGTACAGTTTTTTCATAGCCTTACAGTATCGGCT	2007
Oy	2461	TGTACTAGAAATATAAACACATTTTATTTTGGTTAAAAAATTTTTTTTTT 2514	
Db	2008	TGTACTAGAAATATAAACACATTTTATTTTGGTTAAAAAATTTTTTTTTT 2061	
RESULT 12			
AAf63578			
ID AAf63578 standard; cDNA; 1026 BP.			
AAf63578;			
AC			
AT			
CT			
TT			
Human phosphatase NP_060746_h coding sequence.			

KW	Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW	cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW	congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW	Moebius syndrome; Björnstad syndrome; Bannayan Zonana syndrome;
KW	schizophrenia; hamartoma; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200112819-A2.
XX	
PD	22-FEB-2001.
XX	
XX	11-AUG-2000; 2000WO-US222158.
PF	
XX	
PR	13-AUG-1999; 99US-0149005.
XX	
PA	(SUGEN-) SUGEN INC.
XX	
PI	Plowman CD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX	
DR	WPI: 2001-211226/21.
XX	
PT	P-PsDB; AAB73226.
XX	
PT	New protein phosphatase polypeptide for diagnosing and treating
XX	phosphatase related disorders such as cancer, schizophrenia, cardiac
XX	dysfunction and/or vascular disorders
XX	
PS	Claim 1; Fig 4; 138pp; English.

CC The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase coding sequence
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer.
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Björnsted syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.

	Query Match:	30.7%;	Score	779.6;	DB	22;	Length	1026;
	Best Local Similarity	89.7%;	Pred.	No.1.5e-138;				
	Matches	924;	Conservative	0;	Mismatches	4;	Indels	102;
					Gaps			
OY	853	ATGTCGTGCTGGTGGGCACAGCGGGACCCGAGCTCCCGCATCTTCCCCACCTTACTG	912					
Dd	1	ATGTCGTGCTGGTGGGCACAGCGGGACCCGAGCTCCCGCATCTTCCCCACCTTACTG	60					
OY	913	GGCTCAGAGTGGAACGCCAAGCAACTTGGAGAGCTGCAGAGGAAACAGGGTCAACCAATC	972					
Dd	61	GGCTCAGAGTGGAACGCCAAGCAACTTGGAGAGCTGCAGAGGAAACAGGGTCAACCAATC	120					
OY	973	TTCGAACATGGCCCGGGAGATTGACAACCTTACCCTGAGCGCTTCACTACGCACAATGTG	1032					
Dd	121	TTCGAACATGGCCCGGGAGATTGACAACCTTACCCTGAGCGCTTCACTACGCACAATGTG	180					
OY	1033	CGCCTCGGGGATGTGAGSGATGGGGCCAGCGTCCGGCATCTGGAAAGAGACGACCGCTTC	1092					
Dd	181	CGCCTCGGGGATGTGAGSGATGGGGCCAGCGTCCGGCATCTGGAAAGAGACGACCGCTTC	240					
OY	1093	ATTGAGGCTGCAGAGCACAGGACACCGACCAAGCTGTGGTCCACTGCAAGATGGGGCTCAGC	1152					
Dd	241	ATTGAGGCTGCAGAGCACAGGACACCGACCAAGCTGTGGTCCACTGCAAGATGGGGCTCAGC	300					
OY	1153	CGCTCAGCGGCACAGTGTGGCTATGCCATGAAGCAATACGAATGCAAGCTCGAGCAG	1212					
Dd	301	CGCTCAGCGGCACAGTGTGGCTATGCCATGAAGCAATACGAATGCAAGCTCGAGCAG	360					

OY	1213	GCCTGGGCCACGATGTAGAGAGCTCCGGCCCATTCGGCCCGCCCCAACCTGGCTCCTGGC	1272
OY	1213	GCCTGGGCCACGATGTAGAGAGCTCCGGCCCATTCGGCCCGCCCCAACCTGGCTCCTGGC	1272
Db	361	GCCCCGTGGCCAGTGTGAAGAGCTCCGGCCCATTCGGCCCGCCCCAACCTGGCTCCTGGC	420
OY	1273	CAGTGTCAATTACCAGAGGACACTCTGACG-----	1302
Db	421	CAGTGTCAATTACCAGAGGACACTCTGACG-----	1302
OY	1303	-----	1302
Db	481	AAATGGGTGGGGTCTCCCGAGAGCACCCAGCCCTTGAGTCTTACACCATTTCCA	540
OY	1303	-----GCCAAMCCTGTAGGGTGTGGGAGAGAAGAGTTGTAGCATGGAGAGAGC	1354
Db	541	CTCTTCCTCCGCAACAACCTTAGGGTGTGGGAGAGAGAGTTGTAGCATGGAGAGAGC	600
OY	1355	CAGGCAACCCCGAAAAGAGCCCTGGGGGCGACGGGGCACTGTAAACCTCCGAGGGTCA	1414
Db	601	CAGGCAACCCCGAAAAGAGCCCTGGGGGCGACGGGGCACTGTAAACCTCCGAGGGTCA	658
OY	1415	TGAGGTTCATCACTCTTCTTGAGACCTCCTTGGAGCTGAGAGGACCTCACTAGACAG	1474
Db	659	TGAGGTTCATCACTCTTCTTGAGACCTCCTT- GAGACTGGAGAGCACTCAG - AGACAG	716
OY	1475	TGACATCCGAGAGGTCTTCTCTCCACAGAGTCTTCACTGAAGACCTGTGACGCCCT	1534
Db	717	TGACATBCCAGAGGTCTTCTCTCCACAGAGTCTTCACTGAAGACCTGTGACGCCCT	776
OY	1535	CCCACAGCTTGACAGGACCAAAGGAGGCCAGCAGAGTGTGACAGAGGGGGCTCACTGCCCT	1594
Db	777	CCCACAGCTTGACAGGACCAAAGGAGGCCAGCAGAGTGTGACAGAGGGGGCTCACTGCCCT	836
OY	1595	GAATCCCGCCACTCAAGTGGTTACCTCCAGGGCACTGCCGTGGTGGCCAAACCGAGCCA	1654
Db	837	GAATCCCGCCACTCAAGTGGTTACCTCCAGGGCACTGCCGTGGTGGCCAAACCGAGCCA	896
OY	1655	GGCCTTCCAGAGCAGAGAGAGGGGAGGAGGCCAGGAGGACCGCTGCATTTCTC	1714
Db	897	GGCCTTCCAGAGCAGAGAGAGGGGAGGAGGCCAGGAGGACCGCTGCATTTCTC	956
OY	1715	TAGCCCCAGGTTCCGAGAGGTGTGAGACAGGCCAGCGTCATGACAGTGAGAGAGAGG	1774
Db	957	TACCCCCAGGTTCCGMAAGGTGTGAGACAGGCCAGCGTCATGACAGTGAGAGAGAGG	1016
OY	1775	CGAGGCGCTGA 1784	
Db	1017	CGAGGCGCTGA 1026	
RESULT_13			
ID	ABL40803	standard; DNA: 599 BP.	
XX	ABL40803:		
AC	03-JUL-2002	(first entry)	
DT			
XX			
DE		Human MAP kinase phosphatase-like enzyme DNA fragment.	
KW		Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;	
KW		antistaphylococcal; antidiabetic; anorectic; cytostatic; cardiant; human;	
KW		antiParkinsonian; cerebroprotective; neuroprotective; nootropic; gene;	
KW		neuroleptic; anticovulsant; anti-HIV; antiarrhythmic; hypotensive;	
KW		aniallergic; dermatological; vulnerary; gene therapy; ds.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200220732-A2.	
XX			
PD		14-MAR-2002.	
PF		27-AUG-2001, 2001WO-EP09848.	

XX	07-SEP-2000; 2000US-230709P.
XX	(FARB) BAYER AG.
PA	Llou J;
PB	WPI, 2002-339802/37.
DR	
XX	New human mitogen activated protein kinase phosphatase-like enzyme
PT	polypeptide, regulators of which are useful for preventing, treating
PR	allergies including asthma, diabetes, obesity, cancer and
PT	cardiovascular diseases -
PS	Disclosure: Fig 8; 134pp; English.
XX	
CC	The invention relates to a purified human mitogen activated protein (MAP)
CC	kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC	by standard recombinant methodology. The MAP kinase phosphatase-like
CC	enzyme and encoding polynucleotides are useful for screening for
CC	modulators which are used for treating a MAP kinase phosphatase-like
CC	enzyme dysfunction related disease, such as asthma, a central nervous
CC	system disorder, diabetes, obesity, chronic obstructive pulmonary
CC	disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC	to treat allergies including asthma, allergic rhinitis, atopic
CC	dermatitis, and anaphylaxis, central nervous system disorders such as
CC	stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC	disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC	immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC	including myocardial infarction, ischemic diseases of the heart, arterial
CC	and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC	vascular diseases. The enzyme is useful in diagnostic assays for
CC	detecting diseases and abnormalities or susceptibility to diseases or
CC	abnormalities related to the presence of mutations in the encoding
CC	nucleic acid sequences. The present sequence represents the human MAP
CC	kinase phosphatase-like enzyme DNA fragment.
XX	
SO	Sequence 599 BP; 135 A; 185 C; 183 G; 96 T; 0 other;
	Query Match 22.0%; Score 558.6; DB 24; Length 599;
	Best Local Similarity 98.7%; Pred. No.1.le-96;
Matches	595; Conservative 0; Mismatches 4; Indels 4; Gaps 3
OY	922 TGGAAAGCAACGACTTGAGAGCTGCAGAGAACAAGGTGTCACCACATCTTGAACATG 981
DB	1 TGGAACCGCAGCAAACTTGAGAGCTGCGAGAGAACAGGGTCAACCCACATCTTGAACATG 60
OY	982 GCCGGGAGATTGACAACTTCAACCTGAGGCGCTTCACTTACCAATAATGGGCCCTTG 1041
DB	61 GCCCGGGAATTGACAACTTCAACCTGAGGCGCTTCACTTACCAATAATGGGCCCTTG 120
OY	1042 GATGAGGAGTGGCCCCAGCTGCTGGCGCACTGGAAGGAGACGACCGCTTCAATGAGCT 1101
DB	121 GATGAGGAGTGGCCCCAGCTGCTGGCGCACTGGAAGGAGACGACCGCTTCAATGAGCT 180
OY	1102 GCAAAGCACAGAGGCCACCCAGCTGCTGCTCACTGCAGAAATGGGGCTACCGCTACGG 1161
DB	181 GCMAAGAGCAAGAGGACCCACGCTGCTGGCTCACTGCMAAATGGGGCTACCGCTACGG 240
OY	1162 GCCACAGTCTGCGCTTAATCCATGGAAGCATGAAATGCAAGCTGGAGAGAGGCCCTTGCC 1221
DB	241 GCCACAGTCTGCGCTTAATCCATGGAAGCATGAAATGCAAGCTGGAGAGAGGCCCTTGCC 300
OY	1222 CACGTGCAAGAGCTCCGGCCCATGCGCCGCCCAACCTTGCTTCGCGCAGCTGCGAG 1281
DB	301 CACGTGCAAGAGCTCCGGCCCATGCGCCGCCCAACCTTGCTTCGCGCAGCTGCGAG 360
OY	1282 ATCTTCCAGAGCATCTCTGACGGCCGCAACCACTTAGGTTGGTGCGAGGAAGAAGTGTAG 1341
DB	361 ATCTTCCAGAGCATCTCTGACGGCCGCAACCACTTAGGTTGGTGCGAGGAAGAAGTGTAG 420
OY	1342 CATGGAAGAGACCCAGGCGGCCGGAACAAAGAGCTTGCGGCGCACAGGTATAAC 1401

DR P-PSDB: AAM23524.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX

PS Claim 1: Page 210; 1275pp; English.

XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention.

SQ Sequence 717 BP; 143 A; 222 C; 200 G; 152 T; 0 other;

Query Match 14.9%; Score 377.2; DB 22; Length 717;

Best Local Similarity 99.2%; Pred. No. 2,7e-62;
Matches 379; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 174 GAGTCGACTCCAGCAAGGCGAGCTTGGGCTCCGTGGGCTCTCTGGACTGCA 233
1 ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 697 GGGTCGACTCCAGCAAGGCGAGCTTGGGCTCCGTGGGCTCTCTGGACTGCA 638
OY 234 GGATGGAGGGGCAATGATGATGAGAGAGGCGCAGTTTGAGCCACAGAGAGGCCCC 293
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 637 GGATGGAGGGGCAATGATGATGAGAGAGGCGCAGTTTGAGCCACAGAGAGGCCCC 578
OY 294 GAGTGAGGAGAGCTCCAGCGGGACACAGACAGACTTGGGCAAGATCCAGAGTCCCA 353
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 577 GAGTGAGGAGAGCTCCAGCGGGACACAGACAGACTTGGGCAAGATCCAGAGTCCCA 518
OY 354 GAAGCAGAGAGAGAGAGAGGCGACACTGCACTCATGTGTACGTGAGAGGCCGCAAGA 413
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 517 GAAGCAGAGAGAGAGAGGCGACACTGCACTCATGTGTACGTGAGAGGCCGCAAGA 458
OY 414 TGACATCGGCTGGCAGCCAGCTGAGAGCACCCGGGCTCCCGGCTCCGCTACTGCT 473
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 457 TGACATCGGCTGGCAGCCAGCTGAGAGCACCCGGGCTCCCGGCTCCGCTACTGCT 398
OY 474 GGTAGTTTCTACAGAGAGAGAGGTCTGAGCCAGAGATGAGACGCTCTCTGGGCT 533
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 397 GGTAGTTTCTACAGAGAGAGAGGTCTGAGCCAGAGATGAGACGCTCTCTGGGCT 338
OY 534 GGATTTCCTGACAGACAGCTCC 555
||||||||||||||||||||||||||
Db 337 GGATTTCCTGACAGACAGCTTC 316

Search completed: April 11, 2003, 12:31:11
Job time : 1042.1 secs

Db 629 GTCTACTGCAAGAACGGCCGACGCCGCTGGGCCGCCGCTTGACACCGCGTACCTCATGCGG 688

QY 1189 CAGTAGCAATGCAAGCCCGGAGAGAGGCCCGTGGCAGCTGCAGAGACTCGGCCATCGCC 1248
 Db 689 CACCGCGGGCTCAGCCCTGGGCGAAGGCCCTTCCAGATGTTAAAGCGCTCGGCCCGGTACCA 748
 QY 1249 CGCCCCAACCCCTGGCTCTCGGCCAGCTGCAGATGTATACAGAGGCATCCTGACGGCCAG 1308
 Db 749 GAACCGAACC CGGGCTTCTGCTCTCAGCTCCAACTATAGAGAGAGGCCCTCCAGGCCAG 808
 QY 1309 ACTGAGAGGTGTTGGAGGAGAGAAAGGTGATGAGCATGAGAAAGACCGACGCGC 1362
 Db 809 TCTGCTGCTCAGAGGAGAGAGCCCCAGCCCTTATAGGCTTGGGCGCTTGAGGCTTTGAAGC 862

```

RESULT 2
US-09-704-139-3
: Sequence 3, Application US/09704139
: Patent No. 6420153
: GENERAL INFORMATION:
: APPLICANT: Kapeller, Rosanna
: TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
: FILE REFERENCE: 10448-018001
: CURRENT APPLICATION NUMBER: US/09/7704,139
: CURRENT FILING DATE: 2000-11-01
: PRIOR APPLICATION NUMBER: US 60/185,772
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 531
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-704-139-3

```

Query Match	4.1%	Score 103.4	DB 4	Length 531
Best Local Similarity	52.7%	Pred No. 5.3e-14		
Matches 224, Conservative	0	Mismatches 201	Indels 0	Gaps 0

OY	889	CGATTCCTTCCCCACCCCTCAGCTGAGGGCTCAGATGGAACGGCAGCAAACTGGAGAGCTG	948
Db	61	CGCGTCGGCCCTTACTCTTCTCTCGGGAGCGCGGACGCCGGGCGCGGAGAGCACTG	120
OY	949	CAGAGGAACAGGGTCAACCCACATCTTGGAACTGAGCCCGGGAGATTGACAACTTCTACCTT	1008
Db	121	GGCGGCGCGGAGGACAGCTGTGCTGCAAGGTCTCCCGCAGAGAGCCGGCCCGCGCG	180
OY	1009	GAGGGCTTCACCTTACCCCAATGTGGCCCTCTGGGATGAGGAAGTCGGCCAGCTGCTGCGG	1068
Db	181	CCGCGCTGGCGAGGCTGCGCGGTGCCCGTTCGAGACCCGCGCTGAGAGACCTGCTGGCG	240
OY	1069	CATCGAAGAGAGAGACACCCCTTCATTGAGGCTGCAAGAGACAGAGGACCCAGCTGCTG	1128
Db	241	CACCTGGAGGCCACAGTCTGGCGCCGATGAGAGGCCGCGCGTGGCGCGCGGCGCTTGCTTA	300
OY	1129	GTCACCTGCAAGATGGCGCTCAGCCGCGTCAAGCGCCACAGCTCTGGCCCTTAATCCATGAAG	1188
Db	301	GTCATACGCAAGAACGGGCCGACCCGCTCGGCGCTGTCACCGCGTACTCTTCACGG	360
OY	1189	CAGTACCAATGACAGCCCTGGAGACAGGGCCGTGGCCGACGTGCAGAGAGCTCGGGCCATGGCC	1248
Db	361	CACCGCGGGCTCAGCTCGGCGGAAGGCCCTTCAGATGTGTAAGAAGCGCTGCGCCGGTAACA	420
OY	1249	CGCCCCAACCCCTGGCTTCTCGCCGACGTGCAGATCTTACAGAGGCAATCTGACGCCAGA	1308
Db	421	GAACCGAACCCGGGCTTCTGTCTCAGCTCAGAAATATGAGAGAGGCCCTCCAGGCCAG	480
OY	1309	ACCTG 1313	
Db	481	TCTG 485	

RESULT 3
US-08-530-290-11
; Sequence 11, Application US/08530290

```

: Patent No.5958721
:
: GENERAL INFORMATION:
:
: APPLICANT: Marshall, Christopher John
: APPLICANT: Ashworth, Alan
: APPLICANT: Hughes, David Anthony
: TITLE OF INVENTION: Methods for Screening of Substances for
: TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
: NUMBER OF SEQUENCES: 24
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/530.290
: FILING DATE: 14-DEC-1995
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/GB94/00694
: FILING DATE: 31-MAR-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9402573.1
: FILING DATE: 10-FEB-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9307250.2
: FILING DATE: 07-APR-1993
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bastian, Kevin L.
: REGISTRATION NUMBER: 34,774
: REFERENCE/DOCKET NUMBER: 084611-0000000US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
:
: INFORMATION FOR SEQ ID NO: 11:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1238 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: DNA
:
: US-08-530-290-11

```

[illegible]

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTN016
CLONE: 3041794
US-09-045-973-6

Query Match 2.4%; Score 61.6; DB 4; Length 1729;
Best Local Similarity 46.8%; Pred. No. 0.00016;
Matches 193; Conservative 0; Mismatches 219; Indels 0; Gaps 0;
QY 887 CCCGACATCTTCCCCACCTTACCTGGGCTCAGAGTGGAAAGCAGCAACCTGAGGAGC 946
DB 574 CTCAAATCACCCTCTCTCTATCTCTGGGAGAGGAGGAGTGGCTTCAATCGCACCTCC 633
QY 947 TGGAGAGGAACAGGAGTACCATCTGTAACATGGCCGAGATGACACTTAC 1006
DB 634 TCCAGGCTGTGGCAGTACCTGATTTATGCTTACCTGATGATCCCTAATTTCACT 693
QY 1007 CTGAGCGCTTACCTACACATATGCGCTCTGGAGTGAAGAGTGGCCACCTGCTGC 1066
DB 694 GGGCCCAATTTAGTAAGTGAAGTGGCTGTGGCTGACATGGCGCATGGCCCATGGAC 753
QY 1067 CGCAGCTGAAGAGAGCCACCGCTTATTAAGCTGCAAGAGACAGAGGACCCACGCTGC 1126
DB 754 TGTACTTTGACACCGCTGTGACAGATCCAGATGTGTGACAGAGACAGGAGGACCT 813
QY 1127 TGGTCCACTGCAGATGGGGGTGACCGGCTCAGCGGCGCACAGTGTGCTGTATGGCATGA 1186
DB 814 TGGGACCTGTGCTGACGGGGGTGAGCCGCTCAGCCACGCTGTGTATGCTTACCTGATGA 873
QY 1187 AGCAGTACGATGACAGCTGGAGAGGCGCTGCGCCACAGTGCAGAGAGCTCGGCGCATGC 1246
DB 874 AATTCCACAAGTGTGCTGTGAGAGCGTACACAGTGGGTAAGAAAGCCGCGGACGCTGTCA 933
QY 1247 CCGCGCCCAACCTGCTCTCTCGCGCAGCTGAGATCTACAGGAGCATCT 1298
DB 934 TCAGGCCCAAGTGTGCTGTGAGGACACTGATAGACTACAGAGCGGACACT 985

RESULT 7

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14

Query Match 2.4%; Score 60; DB 1; Length 7218;
Best Local Similarity 2.4%; Pred. No. 0.00068;
Matches 9; Conservative 229; Mismatches 144; Indels 0; Gaps 0;
QY 1835 CTGTGAAGAAACCCCTCAGTCTGTGCGCACACATTCCTCTCAGCTCCGCCCATACC 1894
DB 1059 CTTCGATATTT 1118
QY 1895 CGTCACTACAGCTCCTCCACCCCTGTGACTAGAGGCTCAGCTCCACCCCTGTAC 1954
DB 1119 YY 1178
QY 1955 TACAGCTCAGCTCCTACAGCTTAAGTCCAGGCCAGGCTGTGCTGTGCAAGGCTCA 2014
DB 1179 YY 1238
QY 2015 AGACTTCTACAGGAGTGTGAGAGGAGCTGAAGATCTTTGGGGGCAACAGCACCC 2074
DB 1239 YY 1298
QY 2075 TAGTTTCATCTCACTCAGCCTGACACTCAGCTGTGAGGAGGAATGAAGAGAGC 2134
DB 1299 YY 1358
QY 2135 TTCGCTGCAAAAAGGCTCAGCGCTCCACCCCGCCCTCTGACACTCTCTCTCTC 2194


```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
;
US-09-296-715-25

```

```

Query Match      2.3%; Score 57.4; DB 4; Length 991;
Best local Similarity 76.9%; Pred. No. 0.0011;
Matches 70; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 2450 GTATCTGGCTTGTGACTGAGAAATAAACACATTTTCATATTTGTTAAAAA 2509
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 884 GTATGCTATTTTAACTTTAACTAAAAAGTTCAGTTTAAAAA 943
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2510 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 974
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Search completed: April 11, 2003, 23:54:49
Job time : 183.246 secs

```


Db 241 GGGGACATGATGATGACAGAGGCCAGTTCTGAGCCAAAGAGAGGCCCGGAGTGA 300
OY 301 GAGAGCTCCACGGGGACAGACACTTGGGGAAGGATCCAGAGTCCCGAAGACAG 360
Db 301 GAGAGCTCCACGGGGACAGACACTTGGGGAAGGATCCCGAAGACAG 360
OY 361 GAGAGCAGAGGACAGACCTGACACTTCATGATGATGAGCTGCTGAGGCCGAGATGACATC 420
Db 361 GAGAGCAGAGGACAGACCTGACACTTCATGATGATGAGCTGCTGAGGCCGAGATGACATC 420
OY 421 CGGCTGGACCCGACGTCGAGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTT 480
Db 421 CGGCTGGACCCGACGTCGAGGAGGACCCCGGCTCCCGGCTCCGCTACCTGCTGATGTT 480
OY 481 TCTACAGAGAGAGAGAGTCTGAGCCAGGATGAGAGGCTCTCTCGGAGCGATTTTC 540
Db 481 TCTACAGAGAGAGAGAGTCTGAGCCAGGATGAGAGGCTCTCTCGGAGCGATTTTC 540
OY 541 CCGTACAGAGCTCCCGACGCTGACCCCTGGGCTGGCTTTGGCTTTGGATGACACC 600
Db 541 CCGTACAGAGCTCCCGACGCTGACCCCTGGGCTGGCTTTGGCTTTGGATGACACC 600
OY 601 CAGGTACTTAAATGAGAGAGGGGGCTTGAAGTGAAGCTGTGGGCAAAACCGGATC 660
Db 601 CAGGTACTTAAATGAGAGAGGGGGCTTGAAGTGAAGCTGTGGGCAAAACCGGATC 660
OY 661 TTCAAGCCATCTCCATCCAGACCATGTGCTCTCAGAAAGAGAGAGAGATGAGAGAGCG 720
Db 661 TTCAAGCCATCTCCATCCAGACCATGTGCTCTCAGAAAGAGAGAGAGATGAGAGAGCG 720
OY 721 ATCCGCTGAGCTGTGGAAGTGTGATGTAGTGACCTGAGAGAGTGTACTTCCAAA 780
Db 721 ATCCGCTGAGCTGTGGAAGTGTGATGTAGTGACCTGAGAGAGTGTACTTCCAAA 780
OY 781 GAGATCCGCGAGGCTGGAGCTGGAGCTGGGCTGCCCTCCAGAGATACCGGACTTC 840
Db 781 GAGATCCGCGAGGCTGGAGCTGGAGCTGGGCTGCCCTCCAGAGATACCGGACTTC 840
OY 841 ATGACAAACAGATGCTGTGCTGTGTGGACAGAGGGAGCCGACCTCCGATCTTCCCC 900
Db 841 ATGACAAACAGATGCTGTGCTGTGTGGACAGAGGGAGCCGACCTCCGATCTTCCCC 900
OY 901 CACCTTACTGAGCTGAGAGTGAAGTGAAGCAGCAAACTGAGAGAGCTGTCACTGCAAA 960
Db 901 CACCTTACTGAGCTGAGAGTGAAGTGAAGCAGCAAACTGAGAGAGCTGTCACTGCAAA 960
OY 961 GTACACCAATCTTGAACATGAGGAGGAGATGACAACTTCAACCTGAGCGCTTACC 1020
Db 961 GTACACCAATCTTGAACATGAGGAGGAGATGACAACTTCAACCTGAGCGCTTACC 1020
OY 1021 TACCAATGTGGGCTGTGGATGAGAGTGGGCCAGCTGCTGGCGCACTGGAAGAG 1080
Db 1021 TACCAATGTGGGCTGTGGATGAGAGTGGGCCAGCTGCTGGCGCACTGGAAGAG 1080
OY 1081 ACGGACGCTTCAATGAGCTGCAAGAGCAGAGGCAACCCAGTGTCTGCTCACTGCAAG 1140
Db 1081 ACGGACGCTTCAATGAGCTGCAAGAGCAGAGGCAACCCAGTGTCTGCTCACTGCAAG 1140
OY 1141 ATGGGCTCAGCGCTAGAGGAGGAGGAGGAGTGGGCTATGCGATGGAAGAGCTGCAATGC 1200
Db 1141 ATGGGCTCAGCGCTAGAGGAGGAGGAGGAGTGGGCTATGCGATGGAAGAGCTGCAATGC 1200
OY 1201 AGCGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 AGCGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
OY 1261 GCGCTTCTGCGGACGCTGCAAGATTCACAGGAGATCTCTGAGGCGCAGAACTGAGGGTGG 1320
Db 1261 GCGCTTCTGCGGACGCTGCAAGATTCACAGGAGATCTCTGAGGCGCAGAACTGAGGGTGG 1320
OY 1321 TGGGAGAGAGAGGTTGTAGGATGGAAGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 TGGGAGAGAGAGGTTGTAGGATGGAAGAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380

OY 1381 GGCCACGGGGCAGATATAACCTCCGAGGGGTCATGAGGTCCATCAGTCTTGGAGCCC 1440
Db 1381 GGCCACGGGGCAGATATAACCTCCGAGGGGTCATGAGGTCCATCAGTCTTGGAGCCC 1440
OY 1441 TCCCTTGGAGAGTGGAGAGCAGCTCAGTAGACAGTAGACAGTCCAGAGGTCTTCTTCCC 1500
Db 1441 TCCCTTGGAGAGTGGAGAGCAGCTCAGTAGACAGTAGACAGTCCAGAGGTCTTCTTCCC 1500
OY 1501 ACGAGTCTCAGATGAGAGGCTCTGACAGGCTTCCAGAGTTCAGAGGAGGAGGAGGAGGAG 1560
Db 1501 ACGAGTCTCAGATGAGAGGCTCTGACAGGCTTCCAGAGTTCAGAGGAGGAGGAGGAGGAG 1560
OY 1561 GCGAGAGGTGAGACAGAGGGGCTCAGAGCTGAGAGTCCCGGAGTCAAGTGTATACC 1620
Db 1561 GCGAGAGGTGAGACAGAGGGGCTCAGAGCTGAGAGTCCCGGAGTCAAGTGTATACC 1620
OY 1621 TCCAGGAGAGTGGCGGTGTGGCCAAACCGGAGCCCTTCCAGAGAGAGAGAGAGGAGG 1680
Db 1621 TCCAGGAGAGTGGCGGTGTGGCCAAACCGGAGCCCTTCCAGAGAGAGAGAGAGGAGG 1680
OY 1681 AAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 AAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
OY 1741 GACAGGCGAGCGTGCATGACAGTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1741 GACAGGCGAGCGTGCATGACAGTGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
OY 1801 ACGCTTCCCTGACACTGAAGAGAGATCCACACTCTCTTGAAGAAACCCCTCAGCTGT 1860
Db 1801 ACGCTTCCCTGACACTGAAGAGAGATCCACACTCTCTTGAAGAAACCCCTCAGCTGT 1860
OY 1861 GCGGACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 GCGGACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
OY 1921 CTGTACACTAGGAGCTTACCTCCAGCCCTGTACAGAGGCTTACAGAGGCTTAA 1980
Db 1921 CTGTACACTAGGAGCTTACCTCCAGCCCTGTACAGAGGCTTACAGAGGCTTAA 1980
OY 1981 GTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
Db 1981 GTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
OY 2041 GGGAGTGAAGGATACCTTGGGGGCAACAGACCTAGTTCTCACTGCACTGAGGCTG 2100
Db 2041 GGGAGTGAAGGATACCTTGGGGGCAACAGACCTAGTTCTCACTGCACTGAGGCTG 2100
OY 2101 CACACTCAGCTGTGGCAGGAAATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2101 CACACTCAGCTGTGGCAGGAAATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
OY 2161 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Db 2161 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
OY 2221 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
Db 2221 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
OY 2281 CTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Db 2281 CTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
OY 2341 GTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 2341 GTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
OY 2401 CTTTGTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
Db 2401 CTTTGTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460

Oy 2461 TGTACTGAGAAATAAACAACATTTTCATATTGTTAAAAA 2520
| | | | |
Db 2461 TGTACTGAGAAATAAACAACATTTTCATATTGTTAAAAA 2520
| | | | |
Oy 2521 AAAAAAAAAAAAAAAAAA 2540
| | | | |
Db 2521 AAAAAAAAAAAAAAAAAA 2540
| | | | |

RESULT 2

US-09-761-640-1
: Sequence 1, Application US/09761640
: Patent No. US20020137042A1
: GENERAL INFORMATION:
: APPLICANT: WEI, Ming-Hui et al
: TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: C1000964-CIP
: CURRENT APPLICATION NUMBER: US/09/761,640
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2704
: TYPE: DNA
: ORGANISM: Human
US-09-761-640-1

Query Match 88.4%; Score 2246.4; DB 10; Length 2704;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5;

Oy 1 CCTGTCCTGCGGGTCCAGACTGTCCCGGGGTTGAGGGAAGGGCCGTCCCGGTGC 60
| | | | |
Db 8 CCGTGTCTCGCGGTCCAGAGCTGT-CCGCGGGGTTGAGGGAAGGGCCGTCCCGGTGC 66
| | | | |
Oy 61 CAGCCAGAGTGTCCGCGGCTGTCTAGCCCTGTGCACAGTGAAGCCGTTCGCCCCG 120
| | | | |
Db 67 CAGCCAGAGTGTCCGCGGCTGTCTAGCCCTGTGCACAGTGAAGCCGTTCGCCCCG 126
| | | | |
Oy 121 GGCAGCGGCGCTCCAGCGCCGCTGGGGCCCTGGGACACAGGCGGTCCAGGAAGATGCA 180
| | | | |
Db 127 GGCAGCGGCGCTCCAGCGCCGCTGGGGCCCTGGGACACAGGCGGTCCAGGAAGATGCA 186
| | | | |
Oy 181 CTCACGCGAAGGCAAGCTTTGCGGTGCTCCGTGGGGCTGTCTTGGACTGAGATGGA 240
| | | | |
Db 187 CTCACGCGAAGGCAAGCTTTGCGGTGCTCCGTGGGGCTGTCTTGGACTGAGATGGA 246
| | | | |
Oy 241 GGGGCAATGATGATGCAGAGAGGCGAGTTGTGAGCCAAAGAGAGGCCCCGAGTGA 300
| | | | |
Db 247 GGGGCAATGATGATGCAGAGAGGCGAGTTGTGAGCCAAAGAGAGGCCCCGAGTGA 306
| | | | |
Oy 301 GAGGAGCTCCACGGGGACAGACAGACTTTCGGGCAAGATCCAGAGTCCCGAGAAGCAG 360
| | | | |
Db 307 GAGGAGCTCCACGGGGACAGACAGACTTTCGGGCAAGATCCCGAGAAGCAG 366
| | | | |
Oy 361 GAGGAGCAGAGGCGACCTGTGACTCATGTGACAGCTGTCTAGGCGCGAGATGACATC 420
| | | | |
Db 367 GAGGAGCAGAGGCGACCTGTGACTCATGTGACAGCTGTCTAGGCGCGAGATGACATC 426
| | | | |
Oy 421 GCGCTGGGAGCCGAGGAGGAGCCCGGCTTCGCCGGCTCCGGTACTGCTGTGAGTT 480
| | | | |
Db 427 GCGCTGGGAGCCGAGGAGGAGCCCGGCTTCGCCGGCTCCGGTACTGCTGTGAGTT 486
| | | | |
Oy 481 TGTACACAGAGAAGAGAGTGTGAGCCAGAGATGAGAGGTCCTCTGGGCGCTGATTTTC 540
| | | | |
Db 487 TGTACACAGAGAAGAGAGTGTGAGCCAGAGATGAGAGGTCCTCTGGGCGCTGATTTTC 546
| | | | |
Oy 541 CCGTACACAGAGTCCCGCAGCTGCACCTGGGCTGTGTTTCCCTCTGGAGTACACAC 600
| | | | |
Db 547 CCGTACACAGAGTCCCGCAGCTGCACCTGGGCTGTGTTTCCCTCTGGAGTACACAC 606
| | | | |

Oy 601 CAGGTGTACTTAGATGAGAGACGGGGCTTCACGCTGACGTGTGTGGCAAGCCGATC 660
| | | | |
Db 607 CAGGTGTACTTAGATGAGAGAGCGGGGCTTCACGCTGACGTGTGTGGCAAGCCGATC 666
| | | | |
Oy 661 TTCAAGCCCATCTCCATCCAGACATGT----- 688
| | | | |
Db 667 TTCAAGCCCATCTCCATCCAGACATGTGGGCCACACTCCAGGTATTGCACCAAGCATGT 726
| | | | |
Oy 689 ----- 688
| | | | |
Db 727 GAGGAGCTCTAGGAGCGGGCTTGTACCGGGTGGCAGTGCCTCACCCTGAGGCCAGCAC 786
| | | | |
Oy 689 ----- 688
| | | | |
Db 787 TACCAAGAGAGACTGACATCCGACAGACACTGCTCAATGATGGAGCGGCTATGGCCGAC 846
| | | | |
Oy 689 -----GTCCTCAGAACAGAGCAGATG 711
| | | | |
Db 847 CTGAGTCTCTCGGCTCCAGCCCGAGCCCGAGCCTGGCGGGTCTCCAGAACAGAGCAGATG 906
| | | | |
Oy 712 GAGCAGGCGGATCGTGTGAGGCTGTGGAAGTGTGATGATGATGATGATGATGATGATG 771
| | | | |
Db 907 GAGCAGGCGGATCGTGTGAGGCTGTGGAAGTGTGATGATGATGATGATGATGATGATG 966
| | | | |
Oy 772 ACTTCCAAAGAGATCGGCGAGGCTGTGAGCTGCGCCTGTGGGCTCCCTCCAGCAGTAC 831
| | | | |
Db 967 ACTTCCAAAGAGATCGGCGAGGCTGTGAGCTGCGCCTGTGGGCTCCCTCCAGCAGTAC 1026
| | | | |
Oy 832 CGTGACTTCATGACAGACACAGATGCTGCTGTGGGACAGCGGGACCGAGCCTCCGCG 891
| | | | |
Db 1027 CGTGACTTCATGACAGACACAGATGCTGCTGTGGGACAGCGGGACCGAGCCTCCGCG 1086
| | | | |
Oy 892 ATCTTCCCGCACCCTTACTTGGGCTTCAGAGTGGAAACGACAGCAAACTTGAGAGAGCTGCAG 951
| | | | |
Db 1087 ATCTTCCCGCACCCTTACTTGGGCTTCAGAGTGGAAACGACAGCAAACTTGAGAGAGCTGCAG 1146
| | | | |
Oy 952 AGGAACAGGAGTACCCACATCTTGAACATGGCCGGGAGATTGACAACCTTACCTGTAG 1011
| | | | |
Db 1147 AGGAACAGGAGTACCCACATCTTGAACATGGCCGGGAGATTGACAACCTTACCTGTAG 1206
| | | | |
Oy 1012 CGCTTACCTTACACAGATGTGCGCTTGGGATGAGAGATGCGCCAGCTGTGCGCGAC 1071
| | | | |
Db 1207 CGCTTACCTTACACAGATGTGCGCTTGGGATGAGAGATGCGCCAGCTGTGCGCGAC 1266
| | | | |
Oy 1072 TGGAGAGAGACCCACCGCTTACTTGAAGGCTGCAAGAGCAGACAGGCAACCCAGCTGCTGTGC 1131
| | | | |
Db 1267 TGGAGAGAGACCCACCGCTTACTTGAAGGCTGCAAGAGCAGACAGGCAACCCAGCTGCTGTGC 1326
| | | | |
Oy 1132 CACTGCAAGATGGGCGTCAAGCGCTCAGGCGCAGAGTGTGGCTATGSCATGGAAGCAG 1191
| | | | |
Db 1327 CACTGCAAGATGGGCGTCAAGCGCTCAGGCGCAGAGTGTGGCTATGSCATGGAAGCAG 1386
| | | | |
Oy 1192 TACGAATGACGCTTGAGAGAGGCGCTGCGCCACGTGACAGAGCTTCGGGCCATGCGCCGC 1251
| | | | |
Db 1387 TACGAATGACGCTTGAGAGAGGCGCTGCGCCACGTGACAGAGCTTCGGGCCATGCGCCGC 1446
| | | | |
Oy 1252 CCCAACCTCGGCTTCTCGCGCAGCTGCAAGATCTACAGAGGATCTCTGAGCGCCAGAAC 1311
| | | | |
Db 1447 CCCAACCTCGGCTTCTCGCGCAGCTGCAAGATCTACAGAGGATCTCTGAGCGCCAGAAC 1506
| | | | |
Oy 1312 TTAGGGTGTGGGAGAGAAAGTTGTAGGCTGGAAGAGAGCCAGGAGCCCGCAAAAGA 1371
| | | | |
Db 1507 TTAGGGTGTGGGAGAGAAAGTTGTAGGCTGGAAGAGAGCCAGGAGCCCGCAAAAGA 1566
| | | | |
Oy 1372 AGAGCCTGGGCGCACGGGGACAGTATAAACTTCGAGGGGGTCATGAGTCCATCAGTCTT 1431
| | | | |
Db 1567 AGAGCCTGGG--CCAGGGCGACAGTATAAACTTCGAGGGGGTCATGAGTCCATCAGTCTT 1624
| | | | |
Oy 1432 CTGAGAGCCCTCTCTTGGGAGCTGGAAGACACCTCAATACACAGTACATCCAGAGGCT 1491
| | | | |
Db 1625 CTGAGAGCCCTCTT--GGAGCTTGAAGAGCACTTCAG--AGACAGTACATCCAGAGGCT 1682
| | | | |
Oy 1492 TCTCTTCCACAGAGTCTTACATGAAGAGCCTTGCAGGCCCTTCCACAGCTTTCAGAGGA 1551
| | | | |

```
|||||
Db 1683 TCTCTCCACACAGTCTTACATGAAGACCTTGTGAGCCCTTCCACAGCTTGAACAGA 1742
OY 1552 CCAAGGAGAGCCAGCAGGTGAGACAGGGGGGCTCAGCCCTGCTGGAAGTCCCGCAGTTCAG 1611
Db 1743 CCAAGGAGAGCCAGCAGGTGAGACAGGGGGGCTCAGCCCTGCTGGAAGTCCCGCAGTTCAG 1802
OY 1612 TGGTTACCTTCCAGAGGAGTGGCGTGTGGCCAAACCGAGCCAGGCTTTCAGAGAGCAGG 1671
Db 1803 TGGTTACCTTCCAGAGGAGTGGCGTGTGGCCAAACCGAGCCAGGCTTTCAGAGAGCAGG 1862
OY 1672 AGCAGGGGAGAGGGGAGGGGAGGGGAGAGCCCTGCACTTTCTCTTACGCCAGTTTCGGA 1731
Db 1863 AGCAGGGGAGAGGGGAGGGGAGGGGAGAGCCCTGCACTTTCTCTTACGCCAGTTTCGGA 1922
OY 1732 AGGTGTGAGACAGGCGCAGTGCATGACAGTGAAGAGAGGGGAGGGCTGAGCCCTCA 1791
Db 1923 AGGTGTGAGACAGGCGCAGTGCATGACAGTGAAGAGAGGGGAGGGCTGAGCCCTCA 1982
OY 1792 CACATGCGCAGCGTCCCTGACACTGAAGAGATCCAACTCTTGAAGAAACACCTTC 1851
Db 1983 CACATGCGCAGCGTCCCTGACACTGAAGAGATCCAACTCTTGAAGAAACACCTTC 2042
OY 1852 AGCTGTGTGGCCGACACATCTCTCAGCTCCGCCCATACCCGTCACATACAGCCTCAC 1911
Db 2043 AGCTGTGTGGCCGACACATCTCTCAGCTCCGCCCATACCCGTCACATACAGCCTCAC 2102
OY 1912 CTCCACACCTGTCTACAGGAGCTTACCTCCACCTCCACCTGTCTACATACAGCCTCACTA 1971
Db 2103 CTCCACACCTGTCTACAGGAGCTTACCTCCACCTCCACCTGTCTACATACAGCCTCACTA 2162
OY 1972 CAGCCTTAAGTCCAGAGCCCATGTCTGTCTGTCCAAAGGGCTCAAGACTTTCTTAAGTGGGA 2031
Db 2163 CAGCCTTAAGTCCAGAGCCCATGTCTGTCTGTCCAAAGGGCTCAAGACTTTCTTAAGTGGGA 2222
OY 2032 TGGTGTAGAGAGGAGTGAAGTACTTGTGGGAGCAGACAGCAGCCTTAATTCATTCAGTCACT 2091
Db 2223 TGGTGTAGAGAGGAGTGAAGTACTTGTGGGAGCAGACAGCAGCCTTAATTCATTCAGTCACT 2282
OY 2092 CTAGCCCTGACACTACCTGTGGCAGCAGAAAGGAGAGGCTTCCCGTCAAAAAGGG 2151
Db 2283 CTAGCCCTGACACTACCTGTGGCAGCAGAAAGGAGAGGCTTCCCGTCAAAAAGGG 2342
OY 2152 TCACGCTCCACACCCCGCCCTCCCTGTGACACTCTGTCTCTCCAGTTCACTTCTGTG 2211
Db 2343 TCACGCTCCACACCCCGCCCTCCCTGTGACACTCTGTCTCTCCAGTTCACTTCTGTG 2402
OY 2212 AACGAGCAGAGCCAGGACACAGTGGGCGCCCAAGGAGGAGGAGATCTCAGAGCCGAGC 2271
Db 2403 AACGAGCAGAGCCAGGACACAGTGGGCGCCCAAGGAGGAGGAGATCTCAGAGCCGAGC 2462
OY 2272 CGGAGGAGGCTGGAAGGCTGGGAGATCGCTTCCCTCATCCACTCCACCGGTCCAGGTC 2331
Db 2463 CGGAGGAGGCTGGAAGGCTGGGAGATCGCTTCCCTCATCCACTCCACCGGTCCAGGTC 2522
OY 2332 TTGTGCTGTCTCCAGAGACTCTGTGACACACAGCCAGATTCACAGGGCACCAGGCCAGA 2391
Db 2523 TTGTGCTGTCTCCAGAGACTCTGTGACACACAGCCAGATTCACAGGGCACCAGGCCAGA 2582
OY 2392 GATAGTCTCTTTTGTCTTCTGTGSCCTGAGTACAGTTTTCATAGGCTTACAGT 2451
Db 2583 GATAGTCTCTTTTGTCTTCTGTGSCCTGAGTACAGTTTTCATAGGCTTACAGT 2642
OY 2452 ATCTGCTTTGTACTGAGAAATAAACAATTTTCATATTTGTTAAAAA 2511
Db 2643 ATCTGCTTTGTACTGAGAAATAAACAATTTTCATATTTGTTAAAAA 2702
OY 2512 AA 2513
Db 2703 AA 2704
```

RESULT 3

```
US-09-761-640-7
; Sequence 7, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-7
```

Query Match 88.4%; Score 2246.4; DB 10; Length 2704;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 11; Indels 194; Gaps 5;

```
OY 1 CTTGGTCTGGGGGTCAGGAGTGTCCCGGGGTTGAGGGAAGGGGCGCTGCGCGTGC 60
Db 8 CTTGGTCTGGGGGTCAGGAGTGT-CCGCGGGGTTGAGGGAAGGGGCGCTGCGCGTGC 66
OY 61 CAGCCAGGTGCTCGCGGCTGCTCATAGGCGCTGTACAGTGAAGCCGTTGCGCCCG 120
Db 67 CAGCCAGGTGCTCGCGGCGCTGCTCATAGGCGCTGTACAGTGAAGCCGTTGCGCCCG 126
OY 121 GGCAGCGGGGCTCCACGCGCCCTGGGGGCTGGAGCCAGGGGGTCCAGGGAAGTGTGA 180
Db 127 GGCAGCGGGGCTCCACGCGCCCTGGGGGCTGGAGCCAGGGGGTCCAGGGAAGTGTGA 186
OY 181 CTCACGCGAAGCAGAGCTTTGCGGTGCTCGTGGGGCTGTCTCTGGAGTGCAGAGTGA 240
Db 187 CTCACGCGAAGCAGAGCTTTGCGGTGCTCGTGGGGCTGTCTCTGGAGTGCAGAGTGA 246
OY 241 GGGGACATGATGATGACAGAGGCGCACTTGTAGCCACAGAGAGGCCCGAGTAG 300
Db 247 GGGGACATGATGATGACAGAGGCGCACTTGTAGCCACAGAGAGGCCCGAGTAG 306
OY 301 GAGAGCTCCAGGGGAGCCAGACAGCTTCGGGCAAGATGCCAGATGCCAGAGCAG 360
Db 307 GAGAGCTCCAGGGGAGCCAGACAGCTTCGGGCAAGATGCCAGATGCCAGAGCAG 366
OY 361 GAGGAGCAGAGCAGACCTGACCTCATGTATACAGTCTGAGGCGCGAGAGTACATC 420
Db 367 GAGGAGCAGAGCAGACCTGACCTCATGTATACAGTCTGAGGCGCGAGAGTACATC 426
OY 421 CGCCTGGACGCCAGCTGAGGAGCACC CGGCTCCCGGCTCCGCTACCTGCTGTAATT 480
Db 427 CGCCTGGACGCCAGCTGAGGAGCACC CGGCTCCCGGCTCCGCTACCTGCTGTAATT 486
OY 481 TCTACACGAGAAAGAGAAAGTGTGAGCCAGATGAGACGCTCTCTGGGCGTGGATTC 540
Db 487 TCTACACGAGAAAGAGAAAGTGTGAGCCAGATGAGACGCTCTCTGGGCGTGGATTC 546
OY 541 CCTGACAGAGAGTCCCGCAGCTGACCTGGGCGTGGGCTGTCTCCCTCTGAGAGTACAC 600
Db 547 CCTGACAGAGAGTCCCGCAGCTGACCTGGGCGTGGGCTGTCTCCCTCTGAGAGTACAC 606
OY 601 CAGGTACTTTAGATGAGAGCGGGGCTTACAGCTGACGCTGTGTGGCAAAAGCCGATC 660
Db 607 CAGGTACTTTAGATGAGAGCGGGGCTTACAGGTGACGCTGTGTGGCAAAAGCCGATC 666
OY 661 TTCAGCCCATCTCCATCCAGACCATGT----- 688
Db 667 TTCAGCCCATCTCCATCCAGACCATGTGGGCGCACCTCAGGTATTCACCAAGCATGT 726
OY 689 ----- 688
```

```
Db 727 GAGGAGCTTAGGAGCGGCGCTTGTACCGGGTGGCAGTGCCTTCACCTGGGCGAGCCAG 786
Oy 689 ----- 688
Db 787 TACGAGGAGACTGAACCTCCGACAGAGCTGCTCAATGATGAGCGCTTATGGCCGAC 846
Oy 689 -----GCTCCTCAGAAACAGGAGGAGAGATG 711
Db 847 CTGAGTCTCTCGGCGCTCCGAGCGCCGAGCCTGGCGGGTCTCTCAGAAACAGAGAGAGATG 906
Oy 712 GAGCAGGCGGATCCGTGCTGAGCTGTGGAAGTGTGGATGTAGTGAAGCTTGAGAGTGTG 771
Db 907 GAGCAGGCGGATCCGTGCTGAGCTGTGGAAGTGTGGATGTAGTGAAGCTTGAGAGTGTG 966
Oy 772 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGGCTGGGGCTCCCTCCAGAGATGAC 831
Db 967 ACTTCCAAAGAGATCCGCGAGGCTGTGAGCTGCGGCTGGGGCTCCCTCCAGAGATGAC 1026
Oy 832 CGTGACTTCATGAGACAGCAGATGCTGTGCTGTGAGCAGACGGGAGCCGAGCTCCGCG 891
Db 1027 CGTGACTTCATGAGACAGCAGATGCTGTGCTGTGAGCAGACGGGAGCCGAGCTCCGCG 1086
Oy 892 ATCTTCCCGCCACCTCTTACCTGGGCTCAGAGTGAAGCGACGAAACCTTGAGAGAGCTGCAG 951
Db 1087 ATCTTCCCGCCACCTCTTACCTGGGCTCAGAGTGAAGCGACGAAACCTTGAGAGAGCTGCAG 1146
Oy 952 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATGAGAACCTTACCTGAG 1011
Db 1147 AGGAACAGGCTACCCACATCTTGAACATGGCCCGGAGATGAGAACCTTACCTGAG 1206
Oy 1012 CGCTTACCTACACAGATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCGCGAC 1071
Db 1207 CGCTTACCTACACAGATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCGCGAC 1266
Oy 1072 TGGAAAGGAGACCCACCGCTTCAATTAGGCTGCAAGAGCAGAGGCAACCCAGCTGCTGTC 1131
Db 1267 TGGAAAGGAGACCCACCGCTTCAATTAGGCTGCAAGAGCAGAGGCAACCCAGCTGCTGTC 1326
Oy 1132 CACTCAGAGATGGGGGTACAGCGGCTCAGGCGGCACAGTCTGGCTATGCGCATGGAAGCG 1191
Db 1327 CACTCAGAGATGGGGGTACAGCGGCTCAGGCGGCACAGTCTGGCTATGCGCATGGAAGCG 1386
Oy 1192 TACGATGAGAGCTGAGAGAGGCGCTGCGCCACGTCAGAGAGCTCCGCGCCATTCGCGCG 1251
Db 1387 TACGATGAGAGCTGAGAGAGGCGCTGCGCCACGTCAGAGAGCTCCGCGCCATTCGCGCG 1446
Oy 1232 CCCAACCCCTGGCTTCTCTGCGCGAGCTGCAAGATCTACAGAGGCTCTGAGCGCGCAGAAC 1311
Db 1447 CCCAACCCCTGGCTTCTCTGCGCGAGCTGCAAGATCTACAGAGGCTCTGAGCGCGCAGAAC 1506
Oy 1312 TGAAGGTTGGTGGGAGGAGAAAGTTGTAGGCAATGGAAGAGACCGAGCGCCCGAAGA 1371
Db 1507 TGAAGGTTGGTGGGAGGAGAAAGTTGTAGGCAATGGAAGAGACCGAGCGCCCGAAGA 1566
Oy 1372 AGAGCCTGGGGGACAGGGGCAAGTATAAACCTCCGAGGGGTCATGAGTCACTAGTCTT 1431
Db 1567 AGAGCCTGGG--CAAGGGCAGAGTATAAACCTCCGAGGGGTCATGAGTCACTAGTCTT 1624
Oy 1432 CTGAGAGCCTCTCTGGGAGCTGGAAGACACTGATAGACAGTGAATGCCAGAGGCTCT 1491
Db 1625 CTGAGAGCCTCTCT--GGAGCTGGAAGACACTGAG--AGACAGTGAATGCCAGAGGCTCT 1682
Oy 1492 TCTCTTCCAGAGTCTTCAATGAAAGGCTTGCAGCGCTTCCACAGCTTGAAGA 1551
Db 1683 TCTCTTCCAGAGTCTTCAATGAAAGGCTTGCAGCGCTTCCACAGCTTGAAGA 1742
Oy 1552 CCAGAGGAGGCGCAGAGTGTGAGAGGGGGCTCAGCTGCTCCCTGAAGTCCCGCAGTCA 1611
Db 1743 CCAGAGGAGGCGCAGAGTGTGAGAGGGGGCTCAGCTGCTCCCTGAAGTCCCGCAGTCA 1802
Oy 1612 TGGTTAACCTCCAGGAGCAGTGCCTGTGTGGCCAAACCGACCCAGGCTTCCAGAGAGCAG 1671
Db 1803 TGGTTAACCTCCAGGAGCAGTGCCTGTGTGGCCAAACCGACCCAGGCTTCCAGAGAGCAG 1862

Oy 1672 AGCAGGGGCGAGGGGCGAGGGAGAGCCCTGCAATTTCTCTAGGCCAGATTCCGGA 1731
Db 1863 AGCAGGGGCGAGGGGCGAGGGGCGAGGGAGAGCCCTGCAATTTCTCTAGGCCAGATTCCGGA 1922
Oy 1732 AGTGGTGAAGACAGGCGCAGCGTGCATGACAGTGAAGAGAGGGGCGAGGCTGAGCCCTCA 1791
Db 1923 AGTGGTGAAGACAGGCGCAGCGTGCATGACAGTGAAGAGAGGGGCGAGGCTGAGCCCTCA 1982
Oy 1792 CACATGCCACGCTCCCTGACACACTGAAGAGATCCACACTTCCTTTGGGAAGAACCCCTC 1851
Db 1983 CACATGCCACGCTCCCTGACACACTGAAGAGATCCACACTTCCTTTGGGAAGAACCCCTC 2042
Oy 1852 AGCTGTGGCGGCGACACATTCCTCTCAGCTCGGCGCCCATACCCGTCATACAGCTCAG 1911
Db 2043 AGCTGTGGCGGCGACACATTCCTCTCAGCTCGGCGCCCATACCCGTCATACAGCTCAG 2102
Oy 1912 CTCCACACCCCTGTCACTACAGGCGCTCAGCTCCACCCCTGTCACTACAGCTCAGCTCA 1971
Db 2103 CTCCACACCCCTGTCACTACAGGCGCTCAGCTCCACCCCTGTCACTACAGCTCAGCTCA 2162
Oy 1972 CAGCTTAAAGTCCAGAGCCCATGTCTGCTGTCCAAAGGCTCAAGACTTTCTAACTGGGA 2031
Db 2163 CAGCTTAAAGTCCAGAGCCCATGTCTGCTGTCCAAAGGCTCAAGACTTTCTAACTGGGA 2222
Oy 2032 TGTGGTGAAGGAGCTGAAGATGATCTTGGGGGCAACAGCAGCTGATTCATTCGAAT 2091
Db 2223 TGTGGTGAAGGAGCTGAAGATGATCTTGGGGGCAACAGCAGCTGATTCATTCGAAT 2282
Oy 2092 CTAGCCCTGACACACTCACCTGTGGCAGCGGAATGAAGAACAGACTTCCCGTGAAGAGGG 2151
Db 2283 CTAGCCCTGACACACTCACCTGTGGCAGCGGAATGAAGAACAGACTTCCCGTGAAGAGGG 2342
Oy 2152 TCAGGCTTCCACCCCGCGCCCTCCCTGACCTGCTGTCTCTCCAGTTATTCCTGG 2211
Db 2343 TCAGGCTTCCACCCCGCGCCCTCCCTGACCTGCTGTCTCTCCAGTTATTCCTGG 2402
Oy 2212 AACCAAGCGAGGCGAGGCAACAGTGGAGCCCGCAAGGAGAGAGATCTCAGAGCCGAGC 2271
Db 2403 AACCAAGCGAGGCGAGGCAACAGTGGAGCCCGCAAGGAGAGAGATCTCAGAGCCGAGC 2462
Oy 2272 CGCGGGAGGCTGGAAGGCGTGGCAGATCGCTTCCCTCATCAGCTCCACCGGTCAGGTC 2331
Db 2463 CGCGGGAGGCTGGAAGGCGTGGCAGATCGCTTCCCTCATCAGCTCCACCGGTCAGGTC 2522
Oy 2332 TTTGCTGCTGTCCCAAGACTCTGTGAGACCAAGCCAGATCAAGGGACCAAGGCCAGA 2391
Db 2523 TTTGCTGCTGTCCCAAGACTCTGTGAGACCAAGCCAGATCAAGGGACCAAGGCCAGA 2582
Oy 2392 GATAGCTCTCTTTTGTCTTCTGCGCTTCTGGGCTTGGCTAGTCAATTTTCAATAGCTTACAGT 2451
Db 2583 GATAGCTCTCTTTTGTCTTCTGCGCTTCTGGGCTTGGCTAGTCAATTTTCAATAGCTTACAGT 2642
Oy 2452 ATCTGGCTTTTACTGAGAAATTAACAACATTTTCATATTTGTTAAAAA 2511
Db 2643 ATCTGGCTTTTACTGAGAAATTAACAACATTTTCATATTTGTTAAAAA 2702
Oy 2512 AA 2513
Db 2703 AA 2704

RESULT 4
US-09-761-640-2
; Sequence 2, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: CL000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
```

```

:
:
: CURRENT FILING DATE: 2001-01-18
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
:
: LENGTH: 2852
:
: TYPE: DNA
: ORGANISM: Human
: OS-09-761-640-2

```

Query Match	84.5%;	Score 2145.4;	DB 10;	Length 2852;
Best Local Similarity	89.2%;	Pred. No. 0;		
Matches 2493;	Conservative	0;	Mismatches 11;	Indels 291;
				Gaps 5.

QY	33	GGTTGAGGGAAGGGGGCCGTGCGCCGCTCCAGCCAGCCAGCTGCTCGGGCTGGCTCATTGAGC	92
Dp	2	GGTTGAGGGAAGGGGGCCGTGCGCCGCTCCAGCCAGCCAGCTGCTCGGGCTGGCTCATTGAGC	61
QY	93	CCTGCTCACAETGAGCCGTTTGCCCCGGGGACGCGCGCTCCAGCCGCTGGGGCCCTG	152
Dp	62	CCTGCTCACAETGAGCCGTTTGCCCCGGGGACGCGCGCTCCAGCCGCTGGGGCCCTG	121
QY	153	GGACCAGGCGGTCCAGCGAAGAGTGCATCTCAGCGAAGGAGAGGCTTCGGGGCTCCG	212
Dp	122	GGACCAGGCGGTCCAGCGAAGAGTGCATCTCAGCGAAGGAGAGGCTTCGGGGCTCCG	181
QY	213	TGGGGCTGTCTCTGGGACTGCGAGTGGAGGGGACAAATGATGTGCACAGAGGCCAATTC	272
Dp	182	TGGGGCTGTCTCTGGGACTGCGAGTGGAGGGGACAAATGATGTGCACAGAGGCCAATTC	241
QY	273	TGAGCCCAACAGAAAGGCCCGCGAATGAGAGAGAGTCTCAGGGGACCCAGACACTTCGG	332
Dp	242	TGAGCCCAACAGAAAGGCCCGCGAATGAGAGAGTCTCAGGGGACCCAGACACTTCGG	301
QY	333	GCAAGGATCCAGAGATCCCGCCAGAGCAGAGAGACAGGACGACCTCGACCTCATGCT	392
Dp	302	GCAAGGATCCAGAGATCCCGCCAGAGAGACAGGAGACGACCTCGACCTCATGCT	361
QY	393	ACAGCTGCTGAGGCCCGCAGATGACATTCGCCCTGCGACGCCAGCTGGAGCACCCGGCC	452
Dp	362	ACAGCTGCTGAGGCCCGCAGATGACATTCGCCCTGCGACGCCAGCTGGAGCACCCGGCC	421
QY	453	TCCCGGGTCCGCTACCGTGCCTGGATTTCTACACGAGAGAGAGAGTCTGAGGCCAGGA	512
Dp	422	TCCCGGGTCCGCTACCGTGCCTGGATTTCTACACGAGAGAGAGTCTGAGGCCAGGA	481
QY	513	TGAGAGGCTCTCTCTGGGCGGTGATTTCCCTGACAGCAGCTCCGCCAGCTGCACCCCTGGG	572
Dp	482	TGAGAGGCTCTCTCTGGGCGGTGATTTCCCTGACAGCAGCTCCGCCAGCTGCACCCCTGGG	541
QY	573	CCTGCTCTGGCCCTTGGAGATGACACCCAGGTACTTGAATGGAGACGGGGCTTCAG	632
Dp	542	CCTGCTCTGGCCCTTGGAGATGACACCCAGGTACTTGAATGGAGACGGGGCTTCAG	601
QY	633	CGTAGACGTCTGTTGGGGAAAACCCGGATCTTCAMGCCCATCTCCATCCAGACCAATGT	688
Dp	602	CGTAGACGTCTGTTGGGGAAAACCCGGATCTTCAMGCCCATCTCCATCCAGACCAATGTGGGC	661
QY	689	-----	688
Dp	662	CACACTCCAGGATATTGCACCAAGCATGTGAGGACGCTTAAGGACGCGCCTTGATACGGG	721
QY	689	-----	688
Dp	722	TGGCAGTGCCTTCACACTTGCGGCCACCCACACTACAGGAGAGACTGAACTCCGACAGAGCTG	781
QY	689	-----	688
Dp	782	CCTCAATGATGAGACGGCTATGGCCGACCTGAGTCTCTGGGGCTCCAGCGCGGAGGC	841
QY	689	-----GGTCTCAGAAACAGAGAGCAGATGGAGACAGGCCATCCGTCTCTAGCTGTGGAAAGT	743
Dp	842	TGGCGGGTCTCTCAAAACAGAGAGCAGATGGAGACAGGCCATCCGTCTCTAGCTGTGGAAAGT	901

QY	744	GGGATGTCAGTGAACCTGGAGAGTGTCACTCCAAAGATCCCGAGGGCTGTGACT	803
Db	902	GTTGGATGTCACTGACTGACCTGGAGAGTGTCACTTCCAAAGATCCCGAGGGCTGTGACT	961
QY	804	GGCGCTGGGGGCTCCCGCTCCAGCACTACCGTACTTCATCGAACACAGATGCTGCT	863
Db	962	GGCGCTGGGGGCTCCCGCTCCAGCACTACCGTACTTCATCGAACACAGATGCTGCT	1021
QY	864	GGTGGCACAGCGGGGACCGACGCTCCCGCATCTCCCGACCTTCACGGGCTCAGATG	923
Db	1022	GGTGGCACAGCGGGGACCGACGCTCCCGCATCTCCCGACCTTCACGGGCTCAGATG	1081
QY	924	GAACCCAGCAAACTCGAGAGAGTCTCAGAGGAACAGGGGTCAACCCACATCTTGAACATGGC	983
Db	1082	GAACCCAGCAAACTCGAGAGAGTCTCAGAGGAACAGGGGTCAACCCACATCTTGAACATGGC	1141
QY	984	CCGGCAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAACAATGTGCGCTTGGGA	1043
Db	1142	CCGGCAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAACAATGTGCGCTTGGGA	1201
QY	1044	TGAGGAGTGGCGCCAGCTGCTGCTCCGACATGGGAAGAGAGCACCGCTTCATTGAGGCTGC	1103
Db	1202	TGAGGAGTGGCGCCAGCTGCTGCTCCGACATGGGAAGAGAGCACCGCTTCATTGAGGCTGC	1261
QY	1104	AAGAGCACAGGGGCACCCAGTGTGTCCTACTGCAAGATGGGGGTGACGCGCTCAGCGGC	1163
Db	1262	AAGAGCACAGGGGCACCCAGTGTGTCCTACTGCAAGATGGGGGTGACGCGCTCAGCGGC	1321
QY	1164	CACACTGTGGGCTTATGTCATGAACACATGACAGTAAGCAAGCCGAGGAGCAAGCCCTGCGCA	1223
Db	1322	CACACTGTGGGCTTATGTCATGAACACATGACAGTAAGCAAGCCGAGGAGCAAGCCCTGCGCA	1381
QY	1224	CGTGGAGAGAGCTCGCGCCCATCGCCGCCCAACCCCTGGCTTCTGCGCCAGACTGACAT	1283
Db	1382	CGTGGAGAGAGCTCGCGCCCATCGCCGCCCAACCCCTGGCTTCTGCGCCAGACTGACAT	1441
QY	1284	CTACCAAGGCACTCTGAGC-----	1302
Db	1442	CTACCAAGGCACTCTGAGGCGCCAGCGCCAGACCATGTCTGGAGAGAGAAAGTGGTGG	1501
QY	1303	-----	1305
Db	1502	GGTCTCCCCAGAGAGACCAACCGCCCTGTAAGTCTCTACACATTTCCACTTCTTCGGCC	1561
QY	1306	AGACCTTAGGGGTGTGGGGAGAGAGAGTGTAGCATGGAAGAGCCAGCGACGCC	1365
Db	1562	AGAACCTTAGGGGTGTGGGGAGAGAGAGTGTAGCATGGAAGAGAGCCAGCGACGCC	1621
QY	1366	GAAGAAGAGGCTTGGGGGACAGGGGACAGTATAACCTCCGAGGGGTATGAGGTCCATC	1425
Db	1622	GAAGAAGAGGCTTGGGGGAC--GGCCAGCATTAACCTCCGAGGGGTATGAGGTCCATC	1679
QY	1426	AGTCTTCTGAGACCTCTCTTGGAGCTGGAGAGCACTCAGTAGACAGATGACATGCCAG	1485
Db	1680	AGTCTTCTGAGACCTCTCTT--GGAGCTGGAGAGCACTCAG--AGACAGTGAACATGCCAG	1737
QY	1486	AGGTTCTCTTCCACAGAGTCTTACATGAAGACCTCTGAGGCCCTTCCACAGCTTG	1545
Db	1738	AGGTTCTCTTCCACAGAGTCTTACATGAAGACCTCTGAGGCCCTTCCACAGCTTG	1797
QY	1546	CAAGGACCAAGGAGGAGCCAGCAAGTGTGAGAGAGGGGCTCAGACTTCCCTGAAAGTCCGCC	1605
Db	1798	CAAGGACCAAGGAGGAGCCAGCAAGTGTGAGAGAGGGGCTCAGACTTCCCTGAAAGTCCGCC	1857
QY	1606	AGTCAAGTGTTAACCTCCAGGGCAGTCCGCTGTGTGGCCAAACCGAGCCAGGCTTTCAGG	1665
Db	1858	AGTCAAGTGTTAACCTCCAGGGCAGTCCGCTGTGTGGCCAAACCGAGCCAGGCTTTCAGG	1917
QY	1666	AGCAGAGAGAGGGGACAGGGGCAAGGGGACAGGACCTCGATTTCCTTACGCCAGAGT	1725
Db	1918	AGCAGAGAGAGGGGACAGGGGCAAGGGGACAGGAGACCTCGATTTCCTTACGCCAGAGT	1977
QY	1726	TCCGGAAGGTGTGAGACAGGCACTGTGCATGACAGTGTGAAGAGGAGCGAGGCTTGAG	1785


```
Db 1978 TCCGGAAGTGTGAGACAGGCCAGCGTGCATGACATGGAGAGAGGCGAGGCGCTGAG 2037
QY 1786 CCCGACACATGCCCCAGGCTCCCTGACACTGAGAGATCCACAACCTCTTGGAGAAAC 1845
Db 2038 CCCGACACATGCCCCAGGCTCCCTGACACTGAGAGATCCACAACCTCTTGGAGAAAC 2097
QY 1846 ACCCTCAGCTGTGTGGCGACACATCTCTCAGCTCCGGCCCATACCGCTCACTACAG 1905
Db 2098 ACCCTCAGCTGTGTGGCGACACATCTCTCAGCTCCGGCCCATACCGCTCACTACAG 2157
QY 1906 CCTCAGCTCCACCCCTGTGCTACATAGGCGCTCACCCTCCACCCCTGTGCTACATAG 1965
Db 2158 CCTCAGCTCCACCCCTGTGCTACATAGGCGCTCACCCTCCACCCCTGTGCTACATAG 2217
QY 1966 CTCTTACAGCTTAAAGTCCAGGCGCCATGTGTGCTGTCCAGAGGCGTCAGACTTTCTAA 2025
Db 2218 CTCTTACAGCTTAAAGTCCAGGCGCCATGTGTGCTGTCCAGAGGCGTCAGACTTTCTAA 2277
QY 2026 CTGGGATGTGTAGAGGAGTGAAGTACCTTTGGGGCAACAGACACCTTATTCATTC 2085
Db 2278 CTGGGATGTGTAGAGGAGTGAAGTACCTTTGGGGCAACAGACACCTTATTCATTC 2337
QY 2086 TCAACTCTAGCCCTGCACACTCAGCTGTGCGACGGAATGAAACAGAGCTTCCCGTCAA 2145
Db 2338 TCAACTCTAGCCCTGCACACTCAGCTGTGCGACGGAATGAAACAGAGCTTCCCGTCAA 2397
QY 2146 AAGGGTCAAGCCCTCCACCCCGCCCGCCCTCCGACACTCTGCTCCCTCCGACCTCAT 2205
Db 2398 AAGGGTCAAGCCCTCCACCCCGCCCGCCCTCCGACACTCTGCTCCCTCCGACCTCAT 2457
QY 2206 TCCGTGAACACGAGCGAGGCAACCAAGTGGCCGCCCAAGAGGAGAGATCCCTCAGGC 2265
Db 2458 TCCGTGAACACGAGCGAGGCAACCAAGTGGCCGCCCAAGAGGAGAGATCCCTCAGGC 2517
QY 2266 CCCAGCCGCGGAGGCTGTGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCCACCGGTC 2335
Db 2518 CCCAGCCGCGGAGGCTGTGAAGGCTGGCAGATCGCTTCCCTCATCCACCTCCACCGGTC 2577
QY 2326 CAGGTCTTGTGCTGCCCCAGACCTCTGTGACACAGGCGAGTACACAGGCGACAG 2385
Db 2578 CAGGTCTTGTGCTGCCCCAGACCTCTGTGACACAGGCGAGTACACAGGCGACAG 2637
QY 2386 GCCAGAGATGCTCTTCTTTTGTCTTGTGCGCTGTGCTGAGTCACTTTTCAATAGCT 2445
Db 2638 GCCAGAGATGCTCTTCTTTTGTCTTGTGCGCTGTGCTGAGTCACTTTTCAATAGCT 2697
QY 2446 TACAGTATGCTGCTTGTACTGAGAAATAAACACATTTTCAATTTGTTAAAAAAA 2505
Db 2698 TACAGTATGCTGCTTGTACTGAGAAATAAACACATTTTCAATTTGTTAAAAAAA 2757
QY 2506 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2540
Db 2758 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2792
```

RESULT 5
US-09-955-732-1

```
; Sequence 1, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Mel, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-955-732-1

Query Match 53.0%; Score 1345.6; DB 9; Length 1980.
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 4; Indels 291; Gaps 5;

```
QY 88 ATGCCCCCTGTGACAGTGTGAGGCGTTCCGCCCGGCGAGCGCGGCTCCACGCCGTGGG 147
Db 1 ATGCCCCCTGTGACAGTGTGAGGCGTTCCGCCCGGCGAGCGCGGCTCCACGCCGTGGG 60
QY 148 CCTGGGACACAGCGGCTCCACGAGAGTGTGACTCCAGCGAGGAGAGCTTTCCGCTG 207
Db 61 CCTGGGACACAGCGGCTCCACGAGAGTGTGACTCCAGCGAGGAGAGCTTTCCGCTG 120
QY 208 CTCCGTGGGGCTGTCTGAGGCTCAGAGTGTGAGGAGCAATGTATGTGACGAGAGGCC 267
Db 121 CTCCGTGGGGCTGTCTGAGGCTCAGAGTGTGAGGAGCAATGTATGTGACGAGAGGCC 180
QY 268 AGTTCTGAGCCACAGAGAAAGGCCCGAGTGTGAGGAGGCTCCAGGGGAGCACAGACAG 327
Db 181 AGTTCTGAGCCACAGAGAAAGGCCCGAGTGTGAGGAGGCTCCAGGGGAGCACAGACAG 240
QY 328 TTGCGGCAAGGATCCAGAGTGTCCAGAGAGAGAGAGAGAGAGACACTTCCACTC 387
Db 241 TTGCGGCAAGGATCCAGAGTGTCCAGAGAGAGAGAGAGAGAGACACTTCCACTC 300
QY 388 ATGTGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGGAGGCGCTGGAGGACCC 447
Db 301 ATGTGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGGAGGCGCTGGAGGACCC 360
QY 448 CGGCTCCCGGCTCCGCTACCTGCTGTGTATTTCTACAGAGAGAGAGAGTCTGAGC 507
Db 361 CGGCTCCCGGCTCCGCTACCTGCTGTGTATTTCTACAGAGAGAGAGAGTCTGAGC 420
QY 508 CAGATATGAGAGGCTGTCTGTGGCGCTGATTTCCCTGTACAGCACTCCCGAGTGCAC 567
Db 421 CAGATATGAGAGGCTGTCTGTGGCGCTGATTTCCCTGTACAGCACTCCCGAGTGCAC 480
QY 568 CTGGGCTGTGCTGTGCCCCCTGTGAGTGTGACACCCAGGAGTGTGATGTGAGAGGGGGC 627
Db 481 CTGGGCTGTGCTGTGCCCCCTGTGAGTGTGACACCCAGGAGTGTGATGTGAGAGGGGGC 540
QY 628 TTTCAGCGTGAAGTGTGTGGGCAAAAGCGGATTTTCAAGCCCATCTCATCCAGACCATG 687
Db 541 TTTCAGCGTGAAGTGTGTGGGCAAAAGCGGATTTTCAAGCCCATCTCATCCAGACCATG 600
QY 688 T----- 688
Db 601 TGGGCACACTCCAGGTATTGCAACAGCATGTGAGGAGCTTAGGCGAGGCGCTTGTGA 660
QY 689 ----- 688
Db 661 CCGGGTGGAGTGTCCCTCAGCTGTGGGCCAGCCACTTACAGAGAGAGACTGCCGAACAG 720
QY 689 ----- 688
Db 721 AGTGCTCAATGATGATGAGACGCTATGCGGAGCTGTGCGGCTCCAGGCGC 780
QY 689 -----GTTCTTGAAGACAGAGACAGATGAGAGAGGCGATCGTGTGAGCTGTG 738
Db 781 GAGCGTGGCGGCTCTCAGAAACAGAGACAGATGAGAGAGGCGGATCCGTGTGAGCTGTG 840
QY 739 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 798
Db 841 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 799 GAGCTGGCGCTGGGGCTCCCGCTCCAGAGTACCTGATCTCATCGAACAACAGATGCTG 858
Db 901 GAGCTGGCGCTGGGGCTCCCGCTCCAGAGTACCTGATCTCATCGAACAACAGATGCTG 960
QY 859 CTGCTGTGTGACAGCGGAGCGAGCTCCGCAATTTTCCCACTGTACCTGGGCTCA 918
Db 961 CTGCTGTGTGACAGCGGAGCGAGCTCCGCAATTTTCCCACTGTACCTGGGCTCA 1020
```


Db	781	GAGCCTGGCGGGTCTCTCAAGAACGAGACGATGAGCGAGATCCGTGTAGTGTGG	840
QY	739	AAAGTGTGGATGTCTACGTAGACCTGGAGAGTGTCACTTTCGAAGAGATCGCAGGCTTG	798
Db	841	AAAGTGTGGATGTCTAGTAGACCTGGAGAGTGTCACTTCCAAAGAGATCGCAGGCTTG	900
QY	799	GAGTGGCCCTGGGGGCTCCCCCTCCAGCAGTACGTACTTCATCTGACACAGATGCTG	858
Db	901	GAGTGGCCCTGGGGGCTCCCCCTCCAGCAGTACGTACTTCATCTGACACAGATGCTG	960
QY	859	CTGCTGTGGGACACAGCGGAGCCGAGCTCCCGCATCTTCCCGCACCTGTACTGGGCTCA	918
Db	961	CTGCTGTGTGGACACAGCGGAGCCGAGCTCCCGCATCTTCCCGCACCTGTACTGGGCTCA	1020
QY	919	GAGTGGAAACGACGACAACTCTGGAGAGCTGCAGAGGAACAGGCTACCCACATCTTGAC	978
Db	1021	GAGTGGAAACGACGACAACTCTGGAGAGCTGCAGAGGAACAGGCTACCCACATCTTGAC	1080
QY	979	ATGGCCCCGGAGATTGACAACTTCTACCTTAGAGGCTTACACTACCAAAATGTGGGCTC	1038
Db	1081	ATGGCCCCGGAGATTGACAACTTCTACCTTAGAGGCTTACACTACCAAAATGTGGGCTC	1140
QY	1039	TGGGATGAGAGATGTGGCCCACTGTGCCGCACACTGGAAGAGAGCGACACCCTTCAATTAG	1098
Db	1141	TGGGATGAGAGATGTGGCCCACTGTGCCGCACACTGGAAGAGAGCGACACCCTTCAATTAG	1200
QY	1099	GCTGCAGAGGACACAGGAGACCCACAGTGTGTCTCACTCAAGATGAGGCGTACGGCGCTCA	1158
Db	1201	GCTGCAGAGGACACAGGAGACCCACAGTGTGTCTCACTCAAGATGAGGCGTACGGCGCTCA	1260
QY	1159	GCGGCCACAGTGTCTGGCTATATGCCATGGAAGCAGTACGAATGACAGCTTGAGCAGGCCCTG	1218
Db	1261	GCGGCCACAGTGTCTGGCTATATGCCATGGAAGCAGTACGAATGACAGCTTGAGCAGGCCCTG	1320
QY	1219	GCGCCAGTGCAGAGAGCTCGGCGCCATCGCCGCCCAACCTGTGGTCTCTCGCCAGCTG	1278
Db	1321	GCGCCAGTGCAGAGAGCTCGGCGCCATCGCCGCCCAACCTGTGGTCTCTCGCCAGCTG	1380
QY	1279	CAGATCTTACAGGGCATCTGTACGGCCAGAACTGTGA	1314
Db	1381	CAGATCTTACAGGGCATCTGTACGGCCAGAACTGTGA	1416

```

RESULT 7
US-09-969-708-398/c
; Sequence 398, Application US/09969708
; Patent No. US20020102532A1
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-70
CURRENT APPLICATION NUMBER: US/09/969, 708
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: US/60/237, 606
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237, 608
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237, 425
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 658
SOFTWARE: PatentIn version 3.0
SEQ ID NO 398
LENGTH: 334
TYPE: DNA
; ORGANISM: Homosapiens
US-09-969-708-398

```

QY	2179	TGCACCTCTCTGCTCTCTCCAGTTTCATTCCTGSAACACAGCCAGGCCAGGCAACCACTGTCG	2238
Db	334	TGCACCTCTCTCTCTCTCCAGTTTCATTCCTGSAACACAGCCAGGCCAGGCAACCACTGTCG	275
QY	2239	CCCCAAAGCGACGAGGATCCTCAGGCCCCAGCCGCGGAGAGCTGGAAGAGCGCTGCACAT	2298
Db	274	CCCCAAAGCGACGAGGATCCTCAGGCCCCAGCCGCGGAGAGCTGGAAGAGCGCTGCACAT	215
QY	2299	CGCTTCCCTCATCCACCTCCACCGGTCGACGCTTTTGCTGCTGCCAGACCTCTCTG	2358
Db	214	CGCTTCCCTCATCCACCTCCACCGGTCGACGCTTTTGCTGCTGCCAGACCTCTCTG	155
QY	2359	ACACGACGCCGATCACAGGAGCACAAGGCGCAGAGATAGTCTCTTTTGTCTCTCTGCG	2418
Db	154	ACACGACGCCGATCACAGGAGCACAAGGCGCAGAGATAGTCTCTTTTGTCTCTCTGCG	95
QY	2419	CTCTGGCTAGTCACTTTTTTCATAGCCTTACAGATCTGGCTTTGTACTGAGAAATAAAC	2478
Db	94	CTCTGGCTAGTCACTTTTTTCATAGCCTTACAGATCTGGCTTTGTACTGAGAAATAAAC	35
QY	2479	ACATTTTCATATTGGTTAAACAAAAA 2512	
Db	34	ACATTTTCATATTGGTTAAACAAAAA 1	

```

RESULT 8
US-09-775-925-1
; Sequence 1, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775.925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-775-925-1

```

Query Match	8.3%	Score 212	DB 10	Length 1949
Best Local Similarity	58.9%	Pred. No. 2.1e-42		
Matches 365	Conservative 0	Mismatches 225	Indels 0	Gaps 0
QY 690	GTTCCTCAGAAACGAGGACATGAGACGGCGATCCGTCGTGAGCGTGTGAAGTGTGGA	749		
DB 406	GGCCCATGTGAAGGGGAAAGGACCGAGCGCCTCATCAAGCAAGCACTCCGAAGCATATGAT	465		
QY 750	TGTCAGTACCTCGAGAGTGTCTATCCAAAGAGATCCCGCAGGCTGTGAGCTCGGCT	809		
DB 466	GAGCCAGGATCTAGAAAAATGTGACTTCCAAAGAGATGCTTAATGAATTGAGAAACAGAT	525		
QY 810	GGGGCTCCCTCCACAGCAGTACCGTACTTCATTCGACAACAGATGCTGCTGGTGGC	869		
DB 526	GAAATTGTAACCTGGAAGGAACCTCAGGAATTTTATGACATGATGATGCTACTTATCTTGGG	585		
QY 870	ACAGGGGAGCCAGGCTCCCGCATCTTCCCCCACTCTACCTGGGCTCAGATGGACGC	929		
DB 586	ACAGATGACAAAGCCCTCCCTTATCTTCGATCATCTTTATCTTCGGCTCGAATGGCAATCC	645		
QY 930	ACCAAACTCGAGAGAGCTCAGAGGAACGGGCTCACCCCATCTTGAACATGCGCCGGGA	989		
DB 646	ATCCAAATCTGGAGGAAGCTCAGGGCTCAGGGCTTATTTCATTTTAAAGTTTACCAGAGA	705		
QY 990	GATTGACAATCTTCAACCTGAGCGCTTCACCTACCAACAATGTGGCCCTCTGGATGAGA	1049		
DB 706	AATCATATATTTTTTCTCGGCTTATTTTCATATATATATACATCCGAGTCTACGATGAGA	765		

OY	1050	G T G G C C A A C T G T G T G C C G A C T G A A A G A G A G A C G C A C C G C T C A T T A G G A G C T G C A A G A G C	1109
Db	766	G A C C A C A G A C C T C C T G C C C A C T G A A T A A G G T A T C A T T T T A T P A A C A A A C C C A C A G	825
OY	1110	A C A G G C A C C C A C G T G C T G C T C C A C T G C A A G A T G G G C G T C A G C C C C T C A G C G G C A C A G T	1169
Db	826	G A A C C A T T T C C A A G C C T G T G T G C A T T T G C A A A A T G G G C G T A G A T C G C T G C G C C T C C A G T	885
OY	1170	G C T G G C C T A T T G C C A T G A A G A G A T A C G A A T T G C A G C T T G A G A G C A G G C C C T C G C C A C A G T G C A	1229
Db	886	C A T A G C C T A T G C C A A T G A A G A A T T G C G T G C C T C T G A A A A A G A C T A T A A C T A T A T G T A A A	945
OY	1230	G G A C C T C C G G G C C A T C C C C C G C C C A A C C C T G G C T T C T G C G C C A C G C T G C A G A T A T T A C A	1289
Db	946	G C A A A G G G G A G A T C A C G G G C C C A A C G G G G C T T A T A G A G C A G C G T C T G A C T A T A G A	1005
OY	1290	G G C A T C C T C T A G C G C C A G A A	1309
Db	1006	A G G C A T C T T G G A T G C A A G C A	1025

```

RESULT 9
US-09-775-925-7
: Sequence 7, Application US/09775925
: Patent No. US20010049358A1
: GENERAL INFORMATION:
: APPLICANT: Luche, Ralf M.
: APPLICANT: Wei, Bo
: TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
: TITLE OF INVENTION: PHOSPHATASES
: FILE REFERENCE: 200125.420
: CURRENT APPLICATION NUMBER: US/09/775,925
: CURRENT FILING DATE: 2001-02-01
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 1052
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-775-925-7

```

Query Match	8.28;	Score 208.4;	DB 10;	Length 1052;
Best Local Similarity	57.18;	Pred. No. 1.2e-41;		
Matches 380; Conservative	0;	Mismatches 286;	Indels 0;	Gaps 0;

Oy	695	CAGAAACGAGACAATATGGACACAGCCATCGTCTGAGCTGTGGAAAGTGTTGCATTGCA	754
Dd	200	CTGAACCTGMAACCAACGAAGAAGCTAATTAAAACCAATAAAGGAGATCATGATGCAGA	259
Oy	755	GTCAGCTGGAGAGTGTCATCTCCAAAGATGCCGCGCTGGAGCGCGCCTCGGGGC	814
Dd	260	AGGATTTGGAGAAATATTACATCCAAAGAGATAGAACADAGTTGGAAATGCAAATGGTG	319
Oy	815	TCCCCCTCCACGACTATCCGTACTCATGACACACAGATGCTGCTGCTGTGCACAGC	874
Dd	320	GCAACTTCGCGGAAATTCAGAGAAATTTATAGACATGAATGATAGATCCTTGGTCAA	379
Oy	875	GGGAGCCAGACCTCCCGGATTTCCCCCACCTCACTGGGGCTCAGATGTGAGAACGACGAA	934
Dd	380	TGGATAGCCCTCACACAGATATTTTGAGATGTGTTCTGGGGCTTCAGATGAATGGCTCCA	439
Oy	935	ACCTGAGAGAGACTCAGACGAGAACAGGGTCCACCATCTTGAACATGGCGCGGAGATTG	994
Dd	440	ACTTAGAGAGACTTACACAAACCGAGGGGTACGGATATCTTTGAATGTCACTCGAGAGATAG	499
Oy	995	ACAACCTTCTACCTGAGCGGCTTCACTACACCAATGTGGCCCTGTGGATGTGAAGTGG	1054
Dd	500	ATAACTTCTCCAGAGGCTCTTTGAGATCATTAACCTGGGGATATATGATGAAGAGCAA	559
Oy	1055	CCCACTGCTCCGCGCACTGGAGAGAGACGACCGCTTCATATGAGCTGCAGACGACAGG	1114
Dd	560	CGGATCTCTCGGCTACTGGGAATGACACTTACAAATATCTCTTAAGCGAAGAAGCAATG	619

QY	1115	GCACCCAGTGTCTGTGCTACTGCAGAAATGGGGCTCAGCCGCTCAGGGCGACAGTGTGG	1174
Db	620	GATCTAAATGTGCTTGTGCCTACTGCAGAAATGGGGTGTGATGCTCTCAGCTTCACCGTGAATGG	679
QY	1175	CCATATGCCATGCAACAGCTAGCAATGCAGCCCTGGAGAGAGGCCCTCGGACCCAGTGCAGAGAC	1234
Db	680	CCATATGCAATGAAGGAMTATGTGGCTGGAAATCTGGAGACGAGCTATGACTATGTGAAGAA	739
QY	1235	TCCGGCCATGTGCCCCCGGCCCAACCTGTGGCTTCTGTGGCCAGCTGTGCAGATCTACCAAGGGA	1294
Db	740	GACGAGACGGTAAACCAAGCCCAACCCAAAGCTTCTATGAGACAACTGGAAGAGTATCAGGGGA	799
QY	1295	TCTCTACAGCCAGACACTGTAGGGGTGTGGGGAGAGAAAGGTTGTAGCATGGAAGAGAC	1354
Db	800	TCTTGTGTGGCAAGTCTCTCTAGGCTGTGATTCTATGGAGGAGGAGCAAGCCCTGGGGAGAGA	859
QY	1355	CAGGCA 1360	
Db	860	AAAGCA 865	

```

RESULT 10
; US-09-775-925-5
; Sequence 5, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ. ID NOS.: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-775-925-5

```

Query Match	8.2%	Score 208.4	DB 10	Length 1711
Best Local Similarity	57.1%	Pred. NO. 1.5e-41		
Matches 380; Conservative	0	Mismatches 286	Indels 0	Gaps 0

[illegible]

[illegible]

```

RESULT 11
US-10-002-600-51
Sequence 51, Application US/10002600
Patent No. US2002013707A1
GENERAL INFORMATION:
APPLICANT: Hopkins, Christopher M.
APPLICANT: Peterson, David P.
APPLICANT: Cocks, Benjamin G.
APPLICANT: Hawkins, Philip R.
TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS
FILE REFERENCE: PA-0042 US
CURRENT APPLICATION NUMBER: US/10/002,600
CURRENT FILING DATE: 2001-10-25
PRIOR APPLICATION NUMBER: 60/243,521
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PERL Program
SEQ ID NO 51
LENGTH: 1665
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Template ID: 235631.2
US-10-002-600-51

```

Query Match	3.9%	Pred. 98.8	DB 12	Length 1685
* Best Local Similarity	53.3%	Pred. No. 1.6e-14		
Matches 208; Conservative	0;	Mismatches 182;	Indels 0;	Gaps 0;

[illegible]

Db 913 CGCGCTGTGGCGTGGAGAGAGCCCTTTACATTCGTTAAGCAGCGCGGGGGGATCATCTCC 972

QY 1252 CCCAACCCCTGGCTTCTCTCTGCGCAGCTGCAG 1281

Db 973 CCCAACTTCAGTTTCATGGGCGACGTCGTCG 1002

```

RESULT 12
US-09-808-701-12
: Sequence 12, Application US/09808701
: Patent No. US20020146757A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Xue, Aidong J.
: APPLICANT: Ren, Feiyun
: APPLICANT: Wang, Dunrui
: APPLICANT: Chen, Rui-hong
: APPLICANT: Drmanac, Radjoe T.
: TITLE OF INVENTION: No. US20020146757A1el Nucleic Acids and
: FILE REFERENCE: 790CIP2D
: CURRENT APPLICATION NUMBER: US/09/808,701
: CURRENT FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: 09/540,217
: PRIOR FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: pc_fl_genes Version 2.0
: SEQ ID NO 12
: LENGTH: 969
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (34)..(765)
US-09-808-701-12

```

Query Match	3.8%	Score	96	DB	10	Length	969
Best Local Similarity	56.4%	Pred. No.	6.2e-14				
Matches 202; Conservative	0	Mismatches	150	Indels	6	Gaps	1

QY	349	CCCGAAGCAGAGGAGCAGGAGGAGCAGCCTGCACCCATGATGATACAGTCTAAGGCG	408
Db	235	CACAAAGCATGCAGGTGATCTGCTCTCAACATCTTCAGGTGATGATCAACCTTCTGCTTGC	294
QY	409	CAGATGATCATTCCGCTGGCAGCCAGCTGAGGAGCACCAGGCTCCCGGATCGCTAC	468
Db	295	GAAAGACAGAAATCAAGCTGGCAGTGGCGCTGAGAGACGCGCTGGGCGAACCGGATCCGGTAC	354
QY	469	CTGCGTGTAGTTTCTTACACGAGAAGAGAAAGTCTGACACCAGATGAGACGGTCTCTGTG	528
Db	355	ATGTGTGTGTGTACAGAGAGCGGGCGCCAGGA-----CACCGAGGAATATTTTCTGTG	408
QY	529	GGCGTGGATTTCCCTGCAGCAGCTCCCCACGCTGCACCTGTGGGCTGTGTTCGCCCTC	588
Db	409	GGAGTGGACCTTTCCAGTAAGGAAAGTAAAGCTGCACACATTTGGGAATGTCTCCGACTG	468
QY	589	TGGATGTGACCCACAGGTACTTATGATGAGAAACGGGGGCTCAGCGTGTAGCTGTGGGG	648
Db	469	TGGACGCAACAGAAATCCACTCTGTGATGTGGAATGTGTGGTTCAACGTGAGCAGCAGCAGGA	528
QY	649	CAAAACCGGATCTTAAGCCATCTTCATCTCAGACCATGTGTGCTCTAGAACAGGAGAC	706
Db	529	AGGATGCAACATATTTAAACCTGTGTCTCTCCAGGCGCAATGTGTGTGCGCTCCAGGTGC	586

RESULT 13

```
US-09-865-993-3
: Sequence 3, Application US/09865993
: Publication No. US20030060437A1
: GENERAL INFORMATION:
: APPLICANT: Brett P. Monia
: APPLICANT: Andrew T. Watt
: TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION
: FILE REFERENCE: RTS-0175
: CURRENT APPLICATION NUMBER: US/09/865,993
: CURRENT FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 89
: SEQ ID NO 3
: LENGTH: 2139
: TYPE: DNA
: ORGANISM: Homo sapiens
: NAME/KEY: CDS
: LOCATION: (47)...(1201)
US-09-865-993-3
```

Query Match 3.5%: Score 89.2; DB 9: Length 2139;

Best Local Similarity 51.5%: Pred. No. 4.3e-12; Matches 205; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```
QY 892 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGCTGCAG 951
DB 590 ATCTTCCCCCTCTCTACCTTGAAGTGCCTACCATGCAATGCAAGTTCCTCGCC 649
QY 952 AGGAACAGGTCACCCACATCTGTGAACATGGCCCGGAGATTGAACATCTTACCTTGAG 1011
DB 650 AACTTGACATCAGACCCCTGCTGAATGTCTCCGACGAGGACCTCGAGGCTCGCATGACC 709
QY 1012 CGCTTACCTACACACATGTGGGCTGTGGATGAGAGTGGCCACAGCTGCTGCCGAC 1071
DB 710 CACCTACACTCAAAATGATCCCTTGAAGACACACACAGGCTGCATTAATGCCAC 769
QY 1072 TGAAGAGAGCAGCCGCTTATGAGGCTGCAGAGCAGAGGACCCACGCTGTGCTC 1131
DB 770 TTTCAAGAGCAATAGACTTATGCTGTGTCAGGAGAAAGGAGGACGAGCTCTGCTC 829
QY 1132 CACTGCAGATGGGGCTGAGCGGCTCAGGCGCCACAGTGGGCTATGCCATGAAGCAG 1191
DB 830 CACTGTGAGGCTGGGATCTCCCGTTACCCACCATCTGCAATGGCTTATGAAGACC 889
QY 1192 TAGCAATGAGCCTTGAGCAGAGCCCTGCGCCACGTGAGAGGCTCCGCCATCGCCGC 1251
DB 890 AAGCAGTTCCGCTGAGAGAGGCTTGCATTATCATCAAGACAGAGAGGATGTCTCG 949
QY 1252 CCCAACCTGGCTTCTGCGCCAGCTGAGATCTACCA 1289
DB 950 CCCAATTGGCTTCATGTGGCCAGCTCTGCAGTACGA 987
```

RESULT 14
US-09-865-993-10

: Sequence 10, Application US/09865993
: Publication No. US20030060437A1

: GENERAL INFORMATION:

: APPLICANT: Brett P. Monia

: APPLICANT: Andrew T. Watt

: TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 5 EXPRESSION

: FILE REFERENCE: RTS-0175

: CURRENT APPLICATION NUMBER: US/09/865,993

: CURRENT FILING DATE: 2001-05-25

: NUMBER OF SEQ ID NOS: 89

: SEQ ID NO 10

: LENGTH: 2479

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

US-09-865-993-10

Query Match 3.5%: Score 89.2; DB 9: Length 2479;

Best Local Similarity 51.5%: Pred. No. 4.6e-12; Matches 205; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```
QY 892 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGCTGCAG 951
DB 758 ATCTTCCCCCTCTCTACCTTGAAGTGCCTACCATGCAATGCAAGTTCCTCGCC 817
QY 952 AGGAACAGGTCACCCACATCTGTGAACATGGCCCGGAGATTGAACATCTTACCTTGAG 1011
DB 818 AACTTGACATCAGACCCCTGCTGAATGTCTCCGACGAGGACCTCGAGGCTGTGATGAC 877
QY 1012 CGCTTACCTACACACATGTGGGCTGTGGATGAGAGTGGCCACAGCTGCTGCCGAC 1071
DB 878 CACCTACACTCAAAATGATCCCTTGAAGACAGCCACAGGCTGCATTAATGCTCCAC 937
QY 1072 TGAAGAGAGCAGCCGCTTATGAGGCTGCAGAGCAGAGGACCCACAGCTGTGCTC 1131
DB 938 TTTCAAGAGCAATAGACTTATGCTGTGTCAGGAGAAAGGAGGACGATGCTGCTC 997
QY 1132 CACTGCAGATGGGGCTGAGCGGCTCAGGCGCCACAGTGGGCTATGCCATGAAGCAG 1191
DB 998 CACTGTGAGGCTGGGATCTCCCGTTACCCACCATCTGCAATGCTTACCTTATGAAGACC 1057
QY 1192 TAGCAATGAGCCTTGAGCAGAGCCCTGCGCCACGTGAGAGGCTCCGCCATCGCCGC 1251
DB 1058 AAGCAGTTCCGCTGAGAGAGGCTTGCATTATCATCAAGACAGAGAGGATGTCTCG 1117
QY 1252 CCCAACCTGGCTTCTGCGCCAGCTGAGATCTACCA 1289
DB 1118 CCCAATTGGCTTCATGTGGCCAGCTCTGCAGTACGA 1155
```

RESULT 15
US-09-902-941-1918

: Sequence 1918, Application US/09902941
: Patent No. US20020172952A1

: GENERAL INFORMATION:

: APPLICANT: Henderson, Robert A.

: APPLICANT: Wang, Tonglong

: APPLICANT: Watanabe, Yoshihiro

: APPLICANT: Johnson, Jeffrey C.

: APPLICANT: Rietter, Marc W.

: APPLICANT: Marnerakis, Margarita

: APPLICANT: Carter, Darick

: APPLICANT: Fanger, Gary R.

: APPLICANT: Vedvik, Thomas S.

: APPLICANT: Bangur, Chaitanya S.

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

: FILE REFERENCE: 210121.478C17

: CURRENT APPLICATION NUMBER: US/09/902,941

: CURRENT FILING DATE: 2001-07-10

: NUMBER OF SEQ ID NOS: 2002

: SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 1918

: LENGTH: 1209

: TYPE: DNA

: ORGANISM: Homo sapiens

US-09-902-941-1918

Query Match 3.5%: Score 88; DB 9: Length 1209;

Best Local Similarity 50.5%: Pred. No. 6.5e-12; Matches 214; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

```
QY 892 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGCTGCAG 951
DB 616 ATCTTCCCCCTCTCTACCTGAGAGTGCCTACCATGCTGCCGAGAGACATCTCTGAC 675
QY 952 AGGAACAGGTCACCCACATCTTGAACATGGCCCGGAGATTGAACATCTTACCTTGAG 1011
DB 676 GCCCTGGCATCAGGCTGTGTGAATGTCTCTCGACTGCCCCAACCACTTTGAAGA 735
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2003, 23:58:55 ; Search time 6901.21 Seconds
(Without alignments)
5960.776 Million cell updates/sec

Title: US-09-761-640-3
Perfect score: 2540
Sequence: 1 cctgcctcctgcggctccagc.....aaaaaaaaaaaaaaaaaaaaa 2540

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	851.4	33.5	1004	BM910493
2	795	31.3	1024	BM560192
3	753.8	29.7	772	BQ014446
4	718.2	28.3	1131	BE737389
5	662.8	26.1	943	BM563401
6	659.4	26.0	991	BM471207

7	657	25.9	790	12	BG326709	BG326709	602425578
8	654	25.7	663	12	BG753072	BG753072	602732341
9	642	25.3	653	12	BG679981	BG679981	602626694
10	631	24.8	716	12	BE907514	BE907514	601497473
11	622.8	24.5	932	12	BG340220	BG340220	602438408
12	597.8	23.5	1228	14	BM921421	BM921421	601675476
13	593.2	23.4	890	12	BE901875	BE901875	601675476
14	586.4	23.1	589	9	AL602806	AL602806	DKF2P686L
15	569	22.4	672	12	BG831015	BG831015	602766736
16	562.6	22.1	791	12	BE881145	BE881145	601492016
17	559	22.0	1125	12	BM809903	BM809903	601158526
18	558.6	22.0	599	10	BE280203	BE280203	603628949
19	558	22.0	678	13	BM047611	BM047611	603628949
20	550	21.7	622	12	BE898514	BE898514	601681423
21	548.4	21.6	584	13	BM018531	BM018531	603646365
22	548	21.6	892	12	BF971565	BF971565	602239811
23	537.2	21.1	1061	14	BM906413	BM906413	603040675
24	526.8	20.7	728	13	BI824278	BI824278	603040675
25	517	20.4	937	14	BQ718968	BQ718968	602850749
26	516	20.3	719	10	BE543087	BE543087	601069024
27	508.4	20.0	917	14	BO712114	BO712114	ACENCCOURT
28	498.8	19.6	863	12	BG749256	BG749256	602708085
29	492.8	19.4	496	10	BE348287	BE348287	hw21108.x
30	490.4	19.3	969	10	BE563116	BE563116	601335734
31	482.4	19.0	568	13	BI087083	BI087083	602850749
32	479	18.9	505	13	BM083928	BM083928	imageqc.2
33	477	18.8	574	13	BI007055	BI007055	OV3-RT006
34	473.2	18.6	535	9	AA134926	AA134926	z023d02.S
35	472.4	18.6	846	13	BI820326	BI820326	603036896
36	467.8	18.4	893	12	BG253488	BG253488	602364213
37	464.6	18.3	750	10	BE531347	BE531347	601278540
38	460.8	18.1	860	12	BE127735	BE127735	601810422
39	459.4	18.1	560	13	BI007133	BI007133	OV3-RT006
40	459	18.1	459	9	AI763368	AI763368	w155904.x
41	456.6	18.0	612	13	BM017088	BM017088	603643638
42	455	17.9	962	12	BG166835	BG166835	602345005
43	453	17.8	897	12	BE873337	BE873337	601450290
44	450.8	17.7	790	13	BM008005	BM008005	603617745
45	449	17.7	729	13	BI086720	BI086720	602850066

ALIGNMENTS

RESULT 1
LOCUS BM910493 1004 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6703004 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5453997
ACCESSION BM910493
VERSION BM910493.1 GI:19360872
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1004)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
TISSUE Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LCM1948 row: 9 column: 22
High quality sequence stop: 725.
Location/Qualifiers 1..1004

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:545397"
/clone.lib="NIH_MGC_98"
/tissue.lib="NIH_MGC_98"
/lab_host="astrocytoma grade IV, cell line"
/Note="Organ: brain; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 211 a 299 c 324 g 167 t 3 others

Query Match 33.5%; Score 851.4; DB 14; Length 1004;
Best Local Similarity 97.4%; Pred. No. 3.7e-122;
Matches 875; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

```

OY 159 GGCAGTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGCTTGGCGTCTCCGCGGGC 218
DB 1 GGGGCTCCACGAGAGAGTGCATCCAGCAGAGGAGCAGCTTGGCGTCTCCGCGGGC 60
OY 219 TGTCTGGGACTCCAGGATGGAAGGGACATGATGATGAGCAGAGGCGCAGCTTGAAGC 278
DB 61 TGTCTGGGACTCCAGGATGGAAGGGACATGATGATGAGCAGAGGCGCAGCTTGAAGC 120
OY 279 AACAGAGAGAGGCGCGAGTGAAGAGAGCTCCAGCGGAGCAGACAGACTTCCGCGCAAG 338
DB 121 AACAGAGAGAGGCGCGAGTGAAGAGAGCTCCAGCGGAGCAGACAGACTTCCGCGCAAG 180
OY 339 ATCCAGAGTCCCGCAGAGAGAGAGCAGAGCAGACCTGCACTCTGATGATGATGATGAT 398
DB 181 ATCCAGAGTCCCGCAGAGAGAGAGCAGAGCAGACCTGCACTCTGATGATGATGATGAT 240
OY 399 GCTGAGGCGCAGAGTGCATCCGCTGGAGAGCCAGCTGAGAGCAAGCCGCGCTCCCGC 458
DB 241 GCTGAGGCGCAGAGTGCATCCGCTGGAGAGCCAGCTGAGAGCAAGCCGCGCTCCCGC 300
OY 459 GCTCGGCTACTGCTGATGTTCTACAGAGAGAGAGAGTCTGAGCAGAGATGAGAC 518
DB 301 GCTCGGCTACTGCTGATGTTCTACAGAGAGAGAGAGTCTGAGCAGAGATGAGAC 360
OY 519 GGTCTCTCTGGGCGTGAATTTCCCTGACAGAGCTCCCGAGCTCCAGCTGGCGCTGT 578
DB 361 GGTCTCTCTGGGCGTGAATTTCCCTGACAGAGCTCCCGAGCTCCAGCTGGCGCTGT 420
OY 579 CTGCGGCTCTGGAGTGAACACCCAGGCTGATGATGATGATGATGATGATGATGATGAT 638
DB 421 CTGCGGCTCTGGAGTGAACACCCAGGCTGATGATGATGATGATGATGATGATGATGAT 480
OY 639 GTCTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGTCTGAGA 698
DB 481 GTCTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGTCTGAGA 540
OY 699 ACAGAGAGAGATGAGAGAGGCGGATCGTGTGAGCTGTGAGAAATGTTGATGATGATGAT 758
DB 541 ACAGAGAGAGATGAGAGAGGCGGATCGTGTGAGCTGTGAGAAATGTTGATGATGATGAT 600
OY 759 CTTGAGAGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGAGCTCGCTGGGCTGCC 818
DB 601 CTTGAGAGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGAGCTCGCTGGGCTGCC 660
OY 819 CTTGAGAGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGAGCTCGCTGGGCTGCC 878
DB 661 CTTGAGAGTGTACTTCCAAAGAGATCCGCGAGGCTGTGAGAGCTCGCTGGGCTGCC 720
OY 879 CCGAGCTCCCGGATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGCGAGCAAACT 938
DB 721 CCGAGCTCCCGGATCTTCCCGACCTTACCTGGGCTCAGAGTGAAGCGAGCAAACT 780

```

```

OY 939 GGAGAGCTGCAGAGAGAGAGGCTACCCACATTTGAACAT-GGCCCGGACATTTGACA 997
DB 781 GGAGAGCTGCAGAGAGAGAGGCTACCCACATTTGACATTTGACATTTGACATTTGAC 840
OY 998 ACTTGTACCTTCCGAGGCTTCCATCCATCCAAATGAGGCGCTTGGATGAGAGTGGC 1055
DB 841 ACTTGTACCTTCCGAGGCTTCCATCCATCCAAATGAGGCGCTTGGATGAGAGTGGC 898

```

RESULT 2
BM560192
LOCUS
DEFINITION
AGENCOURT_6597753 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5431968
5', mRNA sequence.
ACCESSION
BM560192
VERSION
BM560192.1 GI:18804414
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1024)
NIH-MGC http://mgi.nci.nih.gov/
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
CONTACT
Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMLN at:
http://image.lml.gov
Plate: LCM1905 row: b column: 01
high quality sequence stop: 655.
location/Qualifiers
1..1024

FEATURES

```

source
1..1024
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5431968"
/clone.lib="NIH_MGC_98"
/tissue.type="astrocytoma grade IV, cell line"
/Note="Organ: brain; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

BASE COUNT 221 a 305 c 328 g 170 t

Query Match 31.3%; Score 795; DB 13; Length 1024;
Best Local Similarity 98.4%; Pred. No. 1.9e-113;
Matches 824; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

```

OY 158 AGCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGCTTGGCGTCTCCGCGGG 217
DB 1 AGCGGCTCCAGCAGAGGAGTGCATCCAGCAGAGGAGCAGCTTGGCGTCTCCGCGGG 60
OY 218 CTGTCCTGGGAGTGTGAGTGAAGGAGATGATGATGATGATGATGATGATGATGATGAT 277
DB 61 CTGTCCTGGGAGTGTGAGTGAAGGAGATGATGATGATGATGATGATGATGATGATGAT 120
OY 278 CAACAGAGAGGCGCGAGTGAAGGAGAGCTCCAGCGGAGCAGACAGACTTGGGCAAG 337
DB 121 CAACAGAGAGGCGCGAGTGAAGGAGAGCTCCAGCGGAGCAGACAGACTTGGGCAAG 180
OY 338 GATCCAGAGTCCCGCAGAGAGAGGAGCAGAGCAGACCTGCACTTATGATGATGATGAT 397
DB 180 GATCCAGAGTCCCGCAGAGAGAGGAGCAGAGCAGACCTGCACTTATGATGATGATGAT

```


QY 2402 TTTTGTCTCTTCTGCGCTCTGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGCTT 2461
|||||
Db 113 TTTTGTCTCTTCTGCGCTCTGCTAGTCAAGTTTTCATAGCCTTACAGTATCTGCTT 54
QY 2462 GTACTGAGAAATAAACACATTTTCATATTGGTTAAAAA 2514
|||||
Db 53 GTACTGAGAAATAAACACATTTTCATATTGGTTAAAAA 1
RESULT 4
BE737389 1131 bp mRNA linear EST 15-SEP-2000
LOCUS 601306112F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640370 5',
DEFINITION mRNA sequence.
ACCESSION BE737389
VERSION BE737389.1 GI:10151381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1131)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Inocyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LMC344 row: h column: 03
High quality sequence stop: 795.
FEATURES
source
1..1131
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640370"
/clone_lib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOH7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald W. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

BASE COUNT 298 a 373 c 271 g 189 t

Query Match 28.3%; Score 718.2; DB 12; Length 1131;
Best Local Similarity 94.4%; Pired. No. 1.3e-101;
Matches 777; Conservative 0; Mismatches 43; Indels 3; Gaps 3;

QY 1684 GGCAGGGGCGAGGAGCCCTGCATTTCTCTACGCCAGCTCCGAGAGCTGTGAGAC 1743
|||||
Db 5 GGGCGGGGCGAGTGAAGAGCTTCATTTCCCTACGCCAGCTCCGAGAGCTGTGAGAC 64
QY 1744 AGGCACAGCGATGACAGTGGAGAGGCGCGAGCCCTGACACATGCGCCAGC 1803
|||||
Db 65 AGGCACAGCGATGACAGTGGAGAGGCGCGAGCCCTGACACATGCGCCAGC 124
QY 1804 CTCCTCTGACACTGAGAGGATCCACAACTCTTGGAGAAAACCCCTACGCTGTGGC 1863
|||||
Db 125 CTCCTCTGACACTGAGAGGATCCACAACTCTTGGAGAAAACCCCTACGCTGTGGC 184
QY 1864 GCACACATTCCTCTACGCTCCGCCCATACCCGCTACAGCTCCACCCCTG 1923
^ ^ ^

Db 185 GCACACATTCCTCTACGCTCCGCCCATACCCGCTACACACCCCTACCTCCACCCCTG 244
QY 1924 TCACATAGGCGCTCACCTCCACCCGCTGACATACAGCCGACCTCTACAGCTTAAGC 1983
Db 245 TCACATAGGCGCTCACCTCCACCCGCTGACATACAGCCGACCTCTCTACAGCTTAAGC 304
QY 1984 CCAGGCCCATGTGTGCTGCTGCTCCAGAGGCTCAAGACTTTCTTAAGTGTGTAGAGG 2043
Db 305 CCAGGCCCATGTGTGCTGCTGCTCCAGAGGCTCAAGACTTTCTTAAGTGTGTAGAGG 364
QY 2044 ACTGAAGTACCTTTGGGGGCAACAGCACCCTAGTTTCTATTTCTACCTTACCTGAC 2103
Db 365 ACTGAAGTACCTTTGGGGGCAACAGCACCCTAGTTTCTATTTCTACCTTACCTGAC 424
QY 2104 ACTCACCTGTGGCAGAGGATGAAGAACAGAGCTTCCCGTCAAAAAAGGTCACGCTCCA 2163
Db 425 ACTCACCTGTGGCAGAGGATGAAGAACAGAGCTTCCCGTCAAAAAAGGTCACGCTCCA 484
QY 2164 CCCCCCCCCCTCCCTGACCTCTGCTCTCTCCAGTTTCATTTCTGAAACCCAGCAGC 2223
Db 485 CCCCCCCCCCTCCCTGACCTCTGCTCTCTCCAGTTTCATTTCTGAAACCCAGCAGC 544
QY 2224 CAGGCAACAGTGGCCCCCAAGGAGGAGGATCTCTAGGCCCCAGCGGGAGGCTG 2283
Db 545 CAGGCAACAGTGGCCCCCAAGGAGGAGGATCTCTAGGCCCCAGCGGGAGGCTG 604
QY 2284 GAAGGCTGGCAGATGCTTCCCTCATCTCACCCTCCAGCTCAGTCTTCTCTCTGTC 2343
Db 605 GAAGGCTGGCAGATGCTTCCCTCATCTCACCCTCCAGCTCAGTCTTCTCTGTC 663
QY 2344 CCCAGACTCTCTGTGACACCCAGATCAGAGGACAGCCAGCCAGAGATGTCTTCT 2403
Db 664 CCAGAG-CTCTGTGTGACACCCAGATCAGAGGCA-CAGGCCAGATGTCTTCTCT 721
QY 2404 TTTGTCTCTTCTGCGCTCTGCTAGTCAATTTTCATACCTTACAGTATCTGCTTGT 2463
Db 722 TTTGTCTCTTCTGCGCTCTGCTAGTCAATTTTCATACCTTACAGTATCTGCTTGT 781
QY 2464 ACTGAGAAATAAACACATTTTCATTTGTTAAAAA 2506
Db 782 ACTGAGAAATAAACACATTTTCATTTGTTAAAAA 824

RESULT 5
BM563401 943 bp mRNA linear EST 20-FEB-2002
LOCUS BM563401
DEFINITION AGENCOURT_6589387 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5441645
5', mRNA sequence.
ACCESSION BM563401
VERSION BM563401.1 GI:18810292
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 943)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LMC1916 row: e column: 06
High quality sequence stop: 618.

FEATURES
source
1..943
Location/Qualifiers

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:5441645"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOT81; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 201 a 284 c 305 g 153 t
ORIGIN

Query Match 26.1%; Score 662.8; DB 13; Length 943;
Best Local Similarity 94.2%; Pred. No. 5e-93;
Matches 710; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

OY 159 GCGGTCACGAGAGAGATCCACTCCAGCAAGCAGCTTGGCGTCCCTGGGGC 218
DB 26 GCGGTCACGAGAGAGATCCACTCCAGCAAGCAGCTTGGCGTCCCTGGGGC 85
OY 219 TGTCCTGGAGCTGAGATGAGAGGAGCAATGATGACAGAGAGCAGTTCGAGCC 278
DB 86 TGTCCTGGAGCTGAGATGAGAGGAGCAATGATGACAGAGAGCAGTTCGAGCC 145
OY 279 AACAGAGAAGGCCCGAGTGAAGAGAGCTCCACGGGAGCAGACAGATTGGGCAAG 338
DB 146 AACAGAGAAGGCCCGAGTGAAGAGAGCTCCACGGGAGCAGACAGATTGGGCAAG 205
OY 339 ATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
DB 206 ATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 265
OY 399 GCTGAGGCGGAGAGATGATCCGCTGGAGCCAGCTGGAGGAGCCCGGCTCCCG 458
DB 266 GCTGAGGCGGAGAGATGATCCGCTGGAGCCAGCTGGAGGAGCCCGGCTCCCG 325
OY 459 GCTCCGCTACTGCTGTGATGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAG 518
DB 326 GCTCCGCTACTGCTGTGATGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAG 385
OY 519 GGTCTCTCTGGGCGGTGATTTCCCTGACAGAGAGCTCCCGAGCTGACCTGG 578
DB 386 GGTCTCTCTGGGCGGTGATTTCCCTGACAGAGAGCTCCCGAGCTGACCTGG 445
OY 579 CTTCGCTCTGAGATGACACCCAGAGTGTACTTGTAGATGAGAGAGAGAGAG 638
DB 446 CTTCGCTCTGAGATGACACCCAGAGTGTACTTGTAGATGAGAGAGAGAGAG 505
OY 639 GTTCGTTGGGCAAAAGCCGATCTTCAGAGCCATCTCATCCAGACATGTGCT 698
DB 506 GTTCGTTGGGCAAAAGCCGATCTTCAGAGCCATCTCATCCAGACATGTGCT 565
OY 699 ACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
DB 566 ACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 625
OY 759 CCGTGAAGATGATCTTCAAAAGAGATCCGACAGCTGTGAGAGTGGGCTCC 817
DB 626 CCGTGAAGATGATCTTCAAAAGAGATCCGACAGCTGTGAGAGTGGGCTCC 685
OY 818 CCTTCACAGAGTACC-GTACTTCATGACAAACAGAGATGCTGCTGGTGACAGGG 876
DB 686 CCTTCACAGAGTACC-GTACTTCATGACAAACAGAGATGCTGCTGGTGACAG 745
OY 877 GACGAGGCTCCCGATCTTCCCTCCACTTACC 910
DB 746 GCGGGGAGCCGAGCCCTCCCTCCCTCTTTC 779

RESULT 6
LOCUS BM471207 991 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6478155 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563111
ACCESSION BM471207
VERSION BM471207.1 GI:18520249
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12293 row: j column: 08
High quality sequence stop: 592.

FEATURES
source
1..991
Location/Qualifiers

/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:5563111"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 199 a 303 c 329 g 159 t 1 others
ORIGIN

Query Match 26.0%; Score 659.4; DB 13; Length 991;
Best Local Similarity 99.0%; Pred. No. 1.6e-92;
Matches 663; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 34 GTTGAAGGAAGGGGCGTCCCGGTGACAGCCAGCTGCTGGGCTGCTCATGGCC 93
DB 1 GTTGAAGGAAGGGGCGTCCCGGTGACAGCCAGCTGCTGGGCTGCTCATGGCC 60
OY 94 CTGTGACAGTGAAGCGCTTCCCGGCGGAGCGGCGCTCCAGCGCCGCTGG 153
DB 61 CTGTGACAGTGAAGCGCTTCCCGGCGGAGCGGCGCTCCAGCGCCGCTGG 120
OY 154 GACCAAGGCGGTCCAGAGAGAGTGAATCCAGAGCAAGAGAGTGTGGGTGCTCGT 213
DB 121 GACCAAGGCGGTCCAGAGAGAGTGAATCCAGAGCAAGAGAGTGTGGGTGCTCGT 180
OY 214 GGGGCTGTCTGGGACTCCAGAGATGAGAGGAGCAATGATGATGACAGAGAGCGAGTTCT 273
DB 181 GGGGCTGTCTGGGACTCCAGAGATGAGAGGAGCAATGATGATGACAGAGAGCGAGTTCT 240
OY 274 GAGCCAAAG 333
DB 241 GAGCCAAAG 300
OY 334 CAAGGATCCAGAGATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
DB 301 CAAGGATCCAGAGATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 394 CAGCTGTGAGAGCGGAGAGATGATCCGCTGGAGCCAGCTGGAGAGAGAGAGAGAG 453

Db 361 CAGCTGCTGAGGGCCGAGATGACATCCGCTGGCAGCCAGCTGGAGGACCCCGCCCT 420
QY 454 CCCGAGCTCCGCTACCTGCTGATGATTTCTACACGAGAGAGAGAGTGTAGCCAGAT 513
Db 421 CCCGAGCTCCGCTACCTGCTGATGATTTCTACACGAGAGAGAGTGTAGCCAGAT 480
QY 514 GAGACGCTCTCTGCGGCTGGATTTCTGACAGCAGCTCCCCAGCTGCACCTGGGC 573
Db 481 GAGACGCTCTCTGCGGCTGGATTTCTGACAGCAGCTCCCCAGCTGCACCTGGGC 540
QY 574 CTGGTCTTGCCCTCTGAGAGTGAACCCAGTACTTATATGAGAGCGGGGCTTACG 633
Db 541 CTGGTCTTGCCCTCTGAGAGTGAACCCAGTACTTATATGAGAGCGGGGCTTACG 600
QY 634 GTGACGCTGTGGTGGCAAGCCGAGTCTTCAGCCCATCTTCATCCAGCATGTGTGCC 693
Db 601 GTGACGCTGTGGTGGCAAGCCGAGTCTTCAGCCCATCTTCATCCAGCATGTGTGCC 660
QY 694 TCAGACAGG 703
Db 661 ACACTCCAGG 670

RESULT 7
BG326709 790 bp mRNA linear EST 27-FEB-2001
LOCUS 602425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
DEFINITION mRNA sequence.
ACCESSION BG326709
VERSION BG326709.1 GI:131333146
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNCM1277 row: k column: 20
High quality sequence stop: 737.

FEATURES
source Location/Qualifiers

1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4563595"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pORF7; site_1: XhoI; site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(s). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 162 a 237 c 268 g 123 t

Query Match 25.9%; Score 657; DB 12; Length 790;
Best Local Similarity 99.2%; Pred. No. 4.2e-92;
Matches 660; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 GGGAAAGGGGCGGTGCCCGGTGCGCAGCCAGTGTCTGCGGGCTGCTCCATGATGCCCTGGT 98
Db 1 GGGAAAGGGGCGGTGCCCGGTGCGCAGCCAGTGTCTGCGGGCTGCTCCATGATGCCCTGGT 60
QY 99 CACAGTGAACCCGTGCCCCCGGCGAGCGCGCTCTCCAGCCCGGTGGGGCCCTGGGAGCA 158
Db 61 CACAGTGAACCCGTGCCCCCGGCGAGCGCGCTCTCCAGCCCGGTGGGGCCCTGGGAGCA 120
QY 159 GCGGTCACAGCAGAGTGTGCACTCCAGCGAAGGCAGAGCTTTGGGTGCTCCGTGGGGC 218
Db 121 GCGGTCACAGCAGAGTGTGCACTCCAGCGAAGGCAGAGCTTTGGGTGCTCCGTGGGGC 180
QY 219 TGTCTGGAGCTGCAAGATGAGAGGGACAAATGATGATGACAGAGAGCCAGTTCTGAGCC 278
Db 181 TGTCTGGAGCTGCAAGATGAGAGGGACAAATGATGATGACAGAGAGCCAGTTCTGAGCC 240
QY 279 AACAGAGAGGCCCCCAGTGAAGAGAGAGTCCACCGGGAGACACAGACTTTGGGGCAAGG 338
Db 241 AACAGAGAGGCCCCCAGTGAAGAGAGAGTCCACCGGGAGACACAGACTTTGGGGCAAGG 300
QY 339 ATCCAGAGTCCCGCAGAGCAGAGAGAGAGAGCCAGCTGACCTGATGTACAGCT 398
Db 301 ATCCAGAGTCCCGCAGAGCAGAGAGAGAGAGCCAGCTGACCTGATGTACAGCT 360
QY 399 GCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGGACACCCGGGCTCCCG 458
Db 361 GCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGGACACCCGGGCTCCCG 420
QY 459 GCTCCGCTACCTGCTGTGTGTCTCTACAGAGAGAGAGAGGCTGTGAGCAGATGAGAC 518
Db 421 GCTCCGCTACCTGCTGTGTGTCTCTACAGAGAGAGAGAGGCTGTGAGCAGATGAGAC 480
QY 519 GGTCTCTGCTGGGCGTGTGATTTCCCTGACAGCAGCTCCCGCAGCTGACCTGGGCTGGT 578
Db 481 GGTCTCTGCTGGGCGTGTGATTTCCCTGACAGCAGCTCCCGCAGCTGACCTGGGCTGGT 540
QY 579 CTTGCCCTCTGTGAGTGAACACCCAGGTGTACTAGATGAGAGCGGGGCTTCAAGCGTAC 638
Db 541 CTTGCCCTCTGTGAGTGAACACCCAGGTGTACTAGATGAGAGCGGGGCTTCAAGCGTAC 600
QY 639 GTTGTGTGGGCAAGCCGATCTTCAAGCCCATCTTCATCCAGACCATGTGTCTCTGAGA 698
Db 601 GTTGTGTGGGCAAGCCGATCTTCAAGCCCATCTTCATCCAGACCATGTGTGTGGCCACT 660
QY 699 ACAGG 703
Db 661 CCAGG 665

RESULT 8
BG753072 663 bp mRNA linear EST 15-MAY-2001
LOCUS 602732241F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875562 5',
DEFINITION mRNA sequence.
ACCESSION BG753072
VERSION BG753072.1 GI:14063725
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 663)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:

http://Image.lnl.gov
Plate: L1CM1757 row: b column: 11
High quality sequence stop: 663.
Location/Qualifiers

FEATURES

SOURCE

1. .663
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4875562"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT 124 a 205 c 231 g 103 t
ORIGIN

Query Match 25.7%; Score 654; DB 12; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.3e-91;
Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 CCGCGGGTTAGGGAAGGGCCCTGCCCCGCTGCCAGCCCAAGTCTCGCGCCCTGCTC 86
DB 10 CCGCGGGTTAGGGAAGGGCCCTGCCCCGCTGCCAGCCCAAGTCTCGCGCCCTGCTC 69
OY 87 CATGCGCCCTGGTACAGAGAGCCGTTCCGCCCGGCGAGCGCGGCTCCACGCCCTGGG 146
DB 70 CATGCGCCCTGGTACAGAGAGCCGTTCCGCCCGGCGAGCGCGGCTCCACGCCCTGGG 129
OY 147 GCCCTGGACCAAGCGGGTCCAGAGAGAGTCCAGCAGCAAGGACGCTTTGGGT 206
DB 130 GCCCTGGACCAAGCGGGTCCAGAGAGAGTCCAGCAGCAAGGACGCTTTGGGT 189
OY 207 GCTCGGTGGGGCTGTCCTGGGACTGAGATGAGAGGGGACATGATGATGACGAGAGGC 266
DB 190 GCTCGGTGGGGCTGTCCTGGGACTGAGATGAGAGGGGACATGATGATGACGAGAGGC 249
OY 267 CAGTTCCTAGAGCCACAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGGACACAGAGA 326
DB 250 CAGTTCCTAGAGCCACAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGGACACAGAGA 309
OY 327 CTTGCGGCAAGGATCCAGAGTCCCGAGAGCAGAGAGCAGACCTTCACCT 386
DB 310 CTTGCGGCAAGGATCCAGAGTCCCGAGAGCAGAGAGCAGACCTTCACCT 369
OY 387 CATGTACAGCTGTGAGCGCGAGAGATGACATCCGCTGGAGCCAGCTGGAGGACCC 446
DB 370 CATGTACAGCTGTGAGCGCGAGAGATGACATCCGCTGGAGCCAGCTGGAGGACCC 429
OY 447 CCGGCTCCCGGCTCCGCTACTGCTGTTGTTCTACAGAGAGAGAGAGTCTGAG 506
DB 430 CCGGCTCCCGGCTCCGCTACTGCTGTTGTTCTACAGAGAGAGAGAGTCTGAG 489
OY 507 CCAGATGAGAGGCTGCTGCTGAGGCTGAGATTTCCCTGACAGACGCTCCCGAGCTGAC 566
DB 490 CCAGATGAGAGGCTGCTGCTGAGGCTGAGATTTCCCTGACAGACGCTCCCGAGCTGAC 549
OY 567 CCGGCGGCTGCTGCTGCTGAGTGAACCCAGGATGATGATGAGAGAGGAGG 626
DB 550 CCGGCGGCTGCTGCTGCTGAGTGAACCCAGGATGATGATGAGAGAGGAGG 609
OY 627 CTTACGCTGAGCTGCTGCTGAGCAAGCCGATCTTCAAGCCCATCTCATCA 680
DB 610 CTTACGCTGAGCTGCTGCTGAGCAAGCCGATCTTCAAGCCCATCTCATCA 663

RESULT 9
Bg679981

LOCUS Bg679981 653 bp mRNA linear EST 01-MAY-2001
DEFINITION 60262694P1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4751758 5',
mRNA sequence.
ACCESSION Bg679981
VERSION Bg679981.1 GI:13911378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 653)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://Image.lnl.gov
Plate: L1CM10608 row: k column: 23
High quality sequence stop: 650.

FEATURES

SOURCE

1. .653
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4751758"
/clone_lib="NCI CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 143 a 223 c 173 g 114 t
ORIGIN

Query Match 25.3%; Score 642; DB 12; Length 653;
Best Local Similarity 99.8%; Pred. No. 9.2e-90;
Matches 653; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1503 GAGTCTTACATGAGACCTCTGACGCTTCCACAGCTTGCAGAGACCAAGGAGGC 1562
DB 1 GAGTCTTACATGAGACCTCTGACGCTTCCACAGCTTGCAGAGACCAAGGAGGC 60
OY 1563 CAGCAGTGAACAGAGGGGCTCAGCCTCGCTGAGTCCCGCAGTCAAGTGGTACCTC 1622
DB 61 CAGCAGTGAACAGAGGGGCTCAGCCTCGCTGAGTCCCGCAGTCAAGTGGTACCTC 120
OY 1623 CAGGCGAGTCCCGTGTGGGCAACCGACCCAGGCTTCCAGAGCAGAGCAGGGGAGC 1682
DB 121 CA-GGCAATGCCGTGTGGGCAACCGACCCAGGCTTCCAGAGCAGAGCAGGGGAGC 179
OY 1683 GGGCAGGGGCGAGGAGACCCCTGCACTTCTCTAGCCCAAGTCCGGAAGGTGGTGA 1742
DB 180 GGGCAGGGGCGAGGAGACCCCTGCACTTCTCTAGCCCAAGTCCGGAAGGTGGTGA 239
OY 1743 CAGGCGAGCTGACATGAGAGTGAAGAGGCGAGGCTGAGCCCTACACATCCAC 1802
DB 240 CAGGCGAGCTGACATGAGAGTGAAGAGGCGAGGCTGAGCCCTACACATCCAC 299
OY 1803 GCTCCCTGACATGAGAGGATCCACAATCTCTTGAAGAAACACCTACGCTGTTGC 1862
DB 300 GCTCCCTGACATGAGAGGATCCACAATCTCTTGAAGAAACACCTACGCTGTTGC 359
OY 1863 CGCACATTTCTCTACACTCCGCCCATATCCGCTACATAGAGCTCACTCCACCCCT 1922
DB 360 CGCACATTTCTCTACACTCCGCCCATATCCGCTACATAGAGCTCACTCCACCCCT 419
OY 1923 GTCACTAGGCGCTACCTCCACCCCTGTCACTACAGCTCACTCAAGCTTAAGT 1982


```

Db 420 GTGACTAGAGGCTACCTCCACCCCTGTCACCTACCTCCTTACGTAACT 479
QY 1993 CCAGAGCCATGTCTGCTGTGTCAGAGGCTCAGACCTTTTAACTGGATGTGTAGAG 2042
Db 480 CCAGAGCCATGTCTGCTGTGTCAGAGGCTCAGACCTTTTAACTGGATGTGTAGAG 539
QY 2043 GACTGAAGGATACCTTTGGGGGCAACACACCTAGTTTCTCAACTACCTTACCTGCA 2102
Db 540 GACTGAAGGATACCTTTGGGGGCAACACACCTAGTTTCTCAACTACCTTACCTGCA 599
QY 2103 CACTCACCCTGGAGAGGATGAAACAGAGCTTCCCTGCAAAAAGGCTCAG 2156
Db 600 CACTCACCCTGGAGAGGATGAAACAGAGCTTCCCTGCAAAAAGGCTCAG 653

```

```

RESULT 10
BE907514 716 bp mRNA linear EST 20-OCT-2000
LOCUS 601497473F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3899418 5',
DEFINITION mRNA sequence.
ACCESSION BE907514
VERSION BE907514.1 GI:10401153
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 716)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM9697 row: a column: 19
High quality sequence stop: 683.
Location/Qualifiers
1..716
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/clone_lib="NIH_MGC_70"
/clone_image="IMAGE:3899418"
/issue_type="epithelioid carcinoma"
/lab_host="PH10B (phage-resistant)"
/site_2="pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
Average insert size 1.1 kb. Library constructed by Life
Technologies."

```

```

BASE COUNT 138 a 213 c 251 g 114 t
ORIGIN
Query Match 24.8% Score 631, DB 12, Length 716,
Best Local Similarity 96.6%, Pred. No. 4.5e-88;
Matches 687; Conservative 0; Mismatches 20; Indels 4; Gaps 4;

```

```

QY 11 CCGGTCACGAGACTGTCGCCGGGGTTGAGGAGAGGCGCGCGGTGGCAGGCCAGT 70
Db 1 CCGGTCACGAGACTGTCGCCGGGGTTGAGGAGAGGCGCGCGGTGGCAGGCCAGT 60
QY 71 GCTCCGCGCTGGCTCCATGGCCCTGTGTACAGTGAACCTTCCGCCCGGAGCGCG 130
Db 61 GCTCCGCGCTGGCTCCATGGCCCTGTGTACAGTGAACCTTCCGCCCGGAGCGCG 119
QY 131 CCTCAGAGCCCGTGGGCGCTGGAGACGAGCGGTGACAGAGAGTGCAGCTCAGCGAA 190
Db 120 CCTCAGAGCCCGTGGGCGCTGGAGACGAGCGGTGACAGAGAGTGCAGCTCAGCGAA 179

```

```

QY 191 GGCACAGCTTTGGGCTGCTCCGTGGGCTGTCTCGGAGCTGCAGGATGAGGGGACAAATG 250
Db 180 GGCACAGCTTTGGGCTGCTCCGTGGGCTGTCTCGGAGCTGCAGGATGAGGGGACAAATG 239
QY 251 ATGATGCGAGAGAGGCGAGTTCTGAGCCACAGAGAAAGGCCCGAGTGAAGAGAGCTCC 310
Db 240 ATGATGCGAGAGAGGCGAGTTCTGAGCCACAGAGAAAGGCCCGAGTGAAGAGAGCTCC 299
QY 311 ACGGGAGACACAGACAGACTTGGGGCAAGGATCCAGAGTCCCGAGAGCAGAGAGCAGA 370
Db 300 ACGGGAGACACAGACAGACTTGGGGCAAGGATCCAGAGTCCCGAGAGCAGAGAGCAGA 359
QY 371 GGCAGACCTGCACTTCATGATGCTACAGTGTGTAGGCGCAGAGATGACATCCGCTGGCAG 430
Db 360 GGCAGACCTGCACTTCATGATGCTACAGTGTGTAGGCGCAGAGATGACATCCGCTGGCAG 419
QY 431 CCCAGCTGAGAGCACCCTGGGCTCCCGGCTCCGCTACCTGCTGTAGTCTTCTACAGAG 490
Db 420 CCCAGCTGAGAGCACCCTGGGCTCCCGGCTCCGCTACCTGCTGTAGTCTTCTACAGAG 479
QY 491 AAGGAGAGGCTGAGACCAAGATGAGAGCGTCTCTGGGCGTGATTTCCCTGACAGCA 550
Db 480 AAGGAGAGGCTGAGACCAAGATGAGAGCGTCTCTGGGCGTGATTTCCCTGACAGCA 539
QY 551 GCTCCCGCAGCTGCACCCCTGGGCTGTCTGGCCCTCTGAGAGTACACCCAGGTGACT 610
Db 540 GCTCCCGCAGCTGCACCCCTGGGCTGTCTGGCCCTCTGAGAGTACACCCAGGTGACT 599
QY 611 TAGATGAGAGACGGGGGCTTACAGCGTGACGTGTGTGGGCAAGCCGATCTCAAGCCCA 670
Db 600 TAGATGAGAGACGGGGGCTTACAGCGTGACGTGTGTGGGCAAGCCGATCTCTCAAG 657
QY 671 TCTCATCCAGACATCTGTGCTCTGAGACAGAGCAGAGCGATGAGGAGAGCCCA 721
Db 658 TCTCATCCAGACATCTGTGCTCTGAGACAGAGCAGAGCGATGAGGAGAGCCCA 707

```

```

RESULT 11
BG340220 932 bp mRNA linear EST 27-FEB-2001
LOCUS BG340220
DEFINITION 602438408F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556401 5',
ACCESSION BG340220
VERSION BG340220.1 GI:13146647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1258 row: p column: 02
High quality sequence stop: 630.
Location/Qualifiers
1..932

```

```

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4556401"
/clone_lib="NIH_MGC_46"
/issue_type="leiomyosarcoma cell line"
/lab_host="PH10B (phage-resistant)"

```

/note="Organ: uterus; Vector: POTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 160 a 284 c 326 g 162 t
ORIGIN

Query Match 24.5%; Score 622.8; DB 12; Length 932;
Best Local Similarity 98.6%; Pred. No. 7.5e-87;
Matches 649; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

OY 46 GGGCTGCCCCGCTGCCAGCCAGGTGCTGGCGCTGCTCCATGGCCCTGTACAGTG 105
DB 1 GGGCGTGGCCGGTGGCAGCCAGGTGCTGGCGCTGCTCCATGGCCCTGTACAGTG 60
OY 106 ACCGCTGCCCCGGGGGAGGGGCGCTCCAGCGCCGCTGGGGGAGCCAGGGGCTC 165
DB 61 ACCGCTTCCCTCCCTGGGAGGGGCGCTCCAGCGCCGCTGGGGAGCCAGGGGCTC 120
OY 166 CAGCAAGAGAGTGAAGTCCAGCGAAGGAGAGCTTTGGCTGCTCCGTTGGGCTGCTCTG 225
DB 121 CAGCAAGAGAGTGAAGTCCAGCGAAGGAGAGCTTTGGCGTCTCCGTTGGGCTGCTCTG 180
OY 226 GGACTGCGAGATGAGAGGGGACATATGATGACAGAGAGGCCAGTTGAGCCAAACAGAG 285
DB 181 GGACTGCGAGATGAGAGGGGACATATGATGACAGAGAGGCCAGTTGAGCCAAACAGAG 240
OY 286 AAGGCCCCGAGTGAAGAGAGAGTCCAGCGGGGAGCCAGACAGATTTGGGGAAAGATCCCA 345
DB 241 AAGGCCCCGAGTGAAGAGAGAGTCCAGCGGGGAGCCAGACAGATTTGGGGAAAGATCCCA 300
OY 346 AGTCCCAAG 405
DB 301 AGTCCCAAG 360
OY 406 CCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
DB 361 CCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
OY 466 TACCTGCTGCTGCTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
DB 420 TACCTGCTGCTGCTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479
OY 526 CTGGGCGTGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
DB 480 CTGGGCGTGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
OY 586 CTCTGAGTGAACCCAGGTGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645
DB 540 CTCTGAGTGAACCCAGGTGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
OY 646 GGGCAAGAGCGGATGCTTCAAGCCCATCTCCATCCAGACATGTGGCTCCCAAGAGAG 703
DB 600 GGGCAAGAGCGGATGCTTCAAG--CCATCTCCATCAAGAGAGATGTGGGCGACACTCCAGG 656

RESULT 12
BM921421 1228 bp mRNA linear EST 12-MAR-2002
LOCUS BM921421
DEFINITION AGENCOURT.6626274 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5752858
5', mRNA sequence.
ACCESSION BM921421
VERSION BM921421.1 GI:19371800
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1228)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov.
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LLM12787 row: 1 column: 11
High quality sequence start: 56
High quality sequence stop: 615.
Location/Qualifiers
1..1228

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752858"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 259 a 390 c 345 g 231 t 3 others
ORIGIN

Query Match 23.5%; Score 597.8; DB 14; Length 1228;
Best Local Similarity 94.7%; Pred. No. 4.9e-83;
Matches 641; Conservative 0; Mismatches 32; Indels 4; Gaps 2;

OY 14 GTCCAGAGACTGTCCGCGGGGTTGAGGAGAGGGGCGGCTCCGCTCCAGCCAGGTGCT 73
DB 50 GGGCGGAAATTCCTGGGAGATGTTGAGGAGAGGGGCGGCTCCGCTCCAGCCAGGTGCT 109
OY 74 CCGGCGCTGGCTCCATGGCCCTGGTACAGTACAGCGCTTCCGCGCGGCGGCGGCT 133
DB 110 CCGGCGCTGGCTCCATGGCCCTGGTACAGTACAGCGCTTCCGCGCGGCGGCGGCT 169
OY 134 CCAGCGCGTGGGCGGCGGCTGGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 193
DB 170 CCAGCGCGTGGGCGGCGGCTGGAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 229
OY 194 AGAGCTTTCGGGTGCTCCGCTGGGCTGTCCCTGGAGCTGCAGAGTGAAGGAGCAATGATG 253
DB 230 AGAGCTTTCGGGTGCTCCGCTGGGCTGTCCCTGGAGCTGCAGAGTGAAGGAGCAATGATG 289
OY 254 ATGCAGCAGAGAGCGCAGTTTGTGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313
DB 290 ATGCAGCAGAGAGCGCAGTTTGTGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349
OY 314 GGGACACAGACAGCTTCCGGGCAAGGATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAG 373
DB 350 GTGACACAGACAGCTTCCGGGCAAGGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAG 409
OY 374 AGACACTGCACCTCATGTGATACAGTCTGAGAGCCGAGAGATGACATCCGCTGGAGAGCC 433
DB 410 AGACACTGCACCTCATGTGATACAGTCTGAGAGCCGAGAGATGACATCCGCTGGAGAGCC 469
OY 434 ACCTGAGAGACACCGGCGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAG 493
DB 470 ACCTGAGAGACACCGGCGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAG 529
OY 494 GAGAAGTCTGAGCCAGATGAG 553

Db	530	GAGAAAGCTGACGACCGAGATGAGAGGGTCTCTCTGGCGCTGGATTTCCCTGACACAGCT	589
OY	554	CCCCAGCTGCACCTGGGGCTGTGTTGGCCCTGTGAGTACACCCAGTGTACTTAG	613
Db	590	CCCCAGCTGCACCTGGGGCTGGGGCTTGGCCCTGTGGAGTACACCCAGGGGTACTTAC	649
OY	614	ATGAGACGGGGGGCTTCACGCT--GACGCTGTGGTGGGAAA--GCCGGATCTTCAAGCCC	669
Db	650	ATGAGAGAGGGGGCTTACACTTGTGAACCTCTGTGGGGCAAAAGCCCGATCTCTAAGGCC	709
OY	670	ATCTTCATCCAGACCAT	686
Db	710	ATCCCCCTCTAAAAAAT	726

RESULT	13
BE901875	
LOCUS	890 bp mRNA linear EST-29-SEP--2000
DEFINITION	G0167547476F1 NIH_MGC_21 Homo sapiens CDNA clone IMAGE:3958314 5' ,
	mRNA sequence.
ACCESSION	BE901875
VERSION	BE901875.1 GI:10391490
KEYWORDS	EST.
SOURCE	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.				
1 (bases 1 to 890)				
NIH-MGC http://mgc.ncl.nih.gov/ .				
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished (1999)				
Contact: Robert Strausberg, Ph.D.				

Email: cs9apbs-r@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.llnl.gov
Plate: L10CM857 row: g, column: 19
High quality sequence stop: 336.

FEATURES	Location/Qualifiers
source	1. .890

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3958314"
/clone_id="NH_MGC_21"
/tissue_type="Choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies). "
```

BASE COUNT	206 a	203 c	214 g	143 t
ORIGIN				
Query Match		23.4%	Score 593.2;	DB 12;
Best Local Similarity		95.1%	Pred. No. 2.8e-82;	Length 890;
Matches 711; Conservative		0;	Mismatches 23;	Indels 14; Gaps 9;

QY	923	60	983
GGAAAGCGACAAACCTGGAGAGTGCAGAGAGAAACAGGGTCAACCATCTTGAACATTGG	982	CCCGGAGAGTTGACAACTTCTACCTTGAGCGGTTTACCTACACCAATGTCGCCCTCTGGG	1042
1	GGAAAGCGACAAACCTGGAGAGTGCAGAGAGAAACAGGGTCAACCATCTTGAACATTGG	60	CCCGGAGAGTTGACAACTTCTACCTTGAGCGGTTTACCTACACCAATGTCGCCCTCTGGG
Db	983	60	983
QY	923 <td>60</td> <td>983</td>	60	983

QY	1043	ATTGAGGATCGGGCCAGCTGCTGGCGGCACCTGGAGGAGAGACGACACCGCTTCATTGAGGGTG	1102
Db	120	ATGAGGATCGGGCCAGCTGCTGGCGGCACCTGGAGGAGAGACGACACCGCTTCATTGAGGGTG	179
QY	1103	CAAGAGCCAGGAGGACCCACGCTGGCTGGTGCACCTGCAGATGGGGGTGACCGCTCAGCGG	1162
Db	180	CAAGAGCCAGGAGGACCCACGCTGGCTGGTGCACCTGCAGATGGGGGTGACCGCTCAGCGG	239
QY	1153	CCACAGTCTGGCCCTATGCTCCATGAAAGCATACGAATGACCTTGGAGCAGGCGCTTGGCC	1222
Db	240	CCACAGTCTGGCCCTATGCTCCATGAAAGCATACGAATGACCTTGGAGCAGGCGCTTGGCC	299
QY	1223	ACGTCACGAGAGCTCCGGGCGCCATCGCGCCGCCCAACCCCTGGGCTTCCCTGGCGCCAGCTGAGA	1282
Db	300	ACGTCACGAGAGCTCCGGGCGCCATCGCGCCGCCCAACCCCTGGGCTTCCCTGGCGCCAGCTGAGA	359
QY	1283	TCTACGAGGGCATCTCGACGGCCAGAAACCTTAGGGGTGGTGGGAGAGAGATGGTGTAGGC	1342
Db	360	TCTACGAGGGCATCTCGACGGCCAGAAACCTTAGGGGTGGTGGGAGAGAGATGGTGTAGGC	418
QY	1343	ATGGAAGAGAGCCAGGACGCGCCCGAAGAAAGAGCTTGGGGCCAGGGGCGACGCTATAAAC	1402
Db	419	ATGGAAGAGAGCCAGGACGCGCCCGAAGAAAGAGCTTGGGGCCAGGGGCGACGCTATAAAC	475
QY	1403	TCCGAGGGGTATGAGGTGCATCACTTCTTGGAGCCCTCCCTTGGGAGCTGGAGAGCAC	1462
Db	476	TCCGAGGGGTATGAGGTGCATCACTTCTTGGAGCCCTCCCTTGGGAGCTGGAGAGCAC	534
QY	1463	TCAGTAGACCAAGTACATATGCAGAGGCTTCTCTTCCACAGAGTCTTCCATGAGAGACC	1522
Db	535	TCAGTAGACCAAGTACATATGCAGAGGCTTCTCTTCCACAGAGTCTTCCATGAGAGACC	593
QY	1523	TCTGAGGCCCTTCCACAGCTTGCAGAGCAACAGGAGGAGCCAGCAGGTGGACAGGGGCC	1582
Db	594	TCTGAGGCCCTTCCACAGCTTGCAGAGCAACAGGAGGAGCCAGCAGGTGGACAGGGGCC	652
QY	1583	TCAGGCTCCCTGGAAGTCCGCCACAGTCAAGTGAACCTCCAGGCGCAGTGGCGTGTGGC	1642
Db	653	TCAGGCTCCCTGGAAGTCCGCCACAGTCAAGTGAACCTCCAGGCGCAGTGGCGTGTGGC	707
QY	1643	CAACCGGACCCAGGCGCTTCCAGAGGACAG	1670
Db	708	CAACCGGACCCAGGCGCTTCCAGAGGACAG	734
RESULT 14			
AL602806			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
MIPS			
Am Klopferstr. 18a D-82152 Martinsried, Germany			
This is the 5' sequence of the clone insert			
Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.			
1 (bases 1 to 589)			
Wambutt, R., Heubner, D., Mewes, W., Weil, B. and Wiemann, S.			
EST (Wambutt, R., Heubner, D., Mewes, W., Weil, B. and Wiemann, S.)			
Unpublished (1999)			
Contact: Wambutt R			
No sl sequence available.			
This clone (DKFP686L2216) is available at the RZPD in Berlin.			
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059			

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```
FEATURES
  source
    Location/Qualifiers
      1. 589
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_id="686 (synonym: hlc3)"
        /tissue_type="human skeletal muscle"
        /dev_stage="adult"
        /lab_host="DH10B"
        /note="Vector: pTRipleX2; Site_1: SfiIA; Site_2: SfiIB;
        cDNA-collection"
BASE COUNT      113 a      181 c      203 g      91 t      1 others
ORIGIN
Query Match      23.1%; Score 586.4; DB 9; Length 589;
Best Local Similarity 99.7%; Pred. No. 3.7e-81;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  71  GCTGCGGCGCTGCTCCTGACAGTGGCCCTGTGACAGTGGAGCCGTTCCGCCCGGCGAGCGCG 130
      |||
Db   1  GCTGCGGCGCTGCTCCTGACAGTGGCCCTGTGACAGTGGAGCCGTTCCGCCCGGCGAGCGCG 60

QY  131 CCTCAGCGCCCGTGGGCGCTGGAGCAAGCGGCTCCAGCGAAGAGTGCATCCAGCGAA 190
      |||
Db   61 CCTCAGCGCCCGTGGGCGCTGGAGCAAGCGGCTCCAGCGAAGAGTGCATCCAGCGAA 120

QY  191 GGCAGAGCTTGGCGGTGCTCCGTGGGGGCTGCTCGGGAGCTGAGAGATGGAGGGGACAATG 250
      |||
Db   121 GGCAGAGCTTGGCGGTGCTCCGTGGGGGCTGCTCGGGAGCTGAGAGATGGAGGGGACAATG 180

QY  251 ATGATGCAAGCAGAGAGCCAGATTCTGAGCCACAGAGAAAGGCCCGAGTAGAGAGAGCTCC 310
      |||
Db   181 ATGATGCAAGCAGAGAGCCAGATTCTGAGCCACAGAGAAAGGCCCGAGTAGAGAGAGCTCC 240

QY  311 ACGGGGACCAACAGACTTCCGGCAAGAGATCCAGAGTCCCGCAGAGCAGAGAGCAGA 370
      |||
Db   241 ACGGGGACCAACAGACTTCCGGCAAGAGATCCAGAGTCCCGCAGAGCAGAGAGCAGA 300

QY  371 GGCAGACCTGACCTCATGTGTACAGTGTGAGGCGCCAGATGACATCCGCTGGCGAG 430
      |||
Db   301 GGCAGACCTGACCTCATGTGTACAGTGTGAGGCGCCAGATGACATCCGCTGGCGAG 360

QY  431 CCCAGTGGAGGACACCCCGGCGCTCCCGGCTCCGCTACCTGCTGATTTCTACACAG 490
      |||
Db   361 CCCAGTGGAGGACACCCCGGCGCTCCCGGCTCCGCTACCTGCTGATTTCTACACAG 420

QY  491 AAGAGAAAGTCTGAGCCAGATGAGACGCTCTCTGGGCGTGGATTCCCTGACAGA 550
      |||
Db   421 AAGAGAAAGTCTGAGCCAGATGAGACGCTCTCTGGGCGTGGATTCCCTGACAGA 480

QY  551 GCTCCCCAGCTGACCTTGCGCTGGCTGCTGCCCCCTGTGAGTAGACCCAGAGTACT 610
      |||
Db   481 GCTCCCCAGCTGACCTTGCGCTGGCTGCTGCCCCCTGTGAGTAGACCCAGAGTACT 540

QY  611 TAGATGAGACGGGGGCTTACGCTGAGCTGTGTTGGGCAAAAGCCGAT 659
      |||
Db   541 TAGATGAGACGGGGGCTTACGCTGAGCTGTGTTGGGCAAAAGCCGAT 589

RESULT 15
LOCUS      BG831015      672 bp      mRNA      linear      EST 22-MAY-2001
DEFINITION 602766736F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908729 5',
            mRNA sequence.
ACCESSION  BG831015.1 GI:14178602
VERSION     BG831015.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 672)
            NIH-MGC http://mgc.nci.nih.gov/.
```

```
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: gcaps-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.lnl.gov
            Plate: LCM1815 row: h column: 10
            High quality sequence stop: 670.

FEATURES
  source
    Location/Qualifiers
      1. 672
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_id="IMAGE:4908729"
        /clone_1id="NIH_MGC_42"
        /tissue_type="epitheloid carcinoma cell line"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: pancreas; Vector: pOT7; Site_1: XhoI;
        Site_2: EcoRI; cDNA made by oligo-dt priming.
        Directionally cloned into EcoRI/XhoI sites using the
        following 5' adaptor: GGCACGAG(G). Size-selected >500bp
        for average insert size 1.8kb. Library constructed by Ling
        Hong in the laboratory of Gerald M. Rubin (University of
        California, Berkeley) using ZAP-cDNA Synthesis kit
        (Stratagene) and Superscript II RT (Life Technologies).
        Note: this is a NIH-MGC Library. |"
BASE COUNT      146 a      219 c      200 g      107 t
ORIGIN
Query Match      22.4%; Score 569; DB 12; Length 672;
Best Local Similarity 98.5%; Pred. No. 1.7e-78;
Matches 606; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY  1301 CGGCCAGAACTGAGAGGTGGTGGGAGAGAGAGGTTGAGGCATGGAAGAGCCAGCA 1360
      |||
Db   61  CGGCCAGAACTGAGAGGTGGTGGGAGAGAGAGGTTGAGGCATGGAAGAGCCAGCA 120

QY  1361 GCCCCGAAAGAGAGCCCTGGGGGCGACGATATTAACCTCCGAGGGTCATGAGT 1420
      |||
Db   121 GCCCCGAAAGAGAGCCCTGGGGGCGACGATATTAACCTCCGAGGGTCATGAGT 178

QY  1421 CCATCAGTCTTGTGAGCCCTCTCTTGGAGCTGGAGACACCTCACTAGACATGACAT 1480
      |||
Db   179 CCATCAGTCTTGTGAGCCCTCTCTTGGAGCTGGAGACACCTCACTAG-AGACCAGTACAT 236

QY  1481 GCCAGAGGCTTCTCTCTCCAGAGGCTTTCACATGAAGACCTGCGACCTCCGCCA 1540
      |||
Db   237 GCCAGAGGCTTCTCTCTCCAGAGGCTTTCACATGAAGACCTTGGACCTTCCCA 296

QY  1541 GCTTGCAAGACCAAGGGAGGCGACAGGTGAGACAGGGGGCTCAGCTGCGCTGAAGTC 1600
      |||
Db   297 GCTTGCAAGACCAAGGGAGGCGACAGGTGAGACAGGGGGCTCAGCTGCGCTGAAGTC 356

QY  1601 CCGCCAGTCAAGTGGTACCTTCAGAGGAGTGGCGTGGTGGCCAAAGCCAGAGGCTT 1660
      |||
Db   357 CCGCCAGTCAAGTGGTACCTTCAGAGGAGTGGCGTGGTGGCCAAAGCCAGAGGCTT 416

QY  1661 CCAGAGCAGAGCAGAGGGGCGAGGGGCGAGGGGAGAGAGCCCTGACATTTCTCTACGCC 1720
      |||
Db   417 CCAGAGCAGAGCAGAGGGGCGAGGGGCGAGGGGAGAGAGCCCTGACATTTCTCTACGCC 476

QY  1721 CAGGTTCCGAAGGTGGTGAAGAGCCAGCGTGCATGACAGTGAAGAGAGGCGAGGC 1780
      |||
Db   477 CAGGTTCCGAAGGTGGTGAAGAGCCAGCGTGCATGACAGTGAAGAGAGGCGAGGC 536

QY  1781 CTGAGCCCTCAGACATGGCCACGCTCCCTTACACACTGAAGAGATCCACATCTCTTGA 1840
      |||
Db   537 CTGAGCCCTCAGACATGGCCACGCTCCCTTACACACTGAAGAGATCCACATCTCTTGA 596
```

Oy	1841	GAACACCCTCAGCTGTCGTGGCGGCACAATTCTCTCAGCTCCGCCCATTAACCCCTTAC	1900
Dd	597	GAAACACCTTACGTTGTTGGCGCACACTTCTCTCAGCTCCGCCCATTAACCCCTTAC	656
Oy	1901	TACAGCCTCACTTCC	1915
Dd	657	TACAGCCTCACTTCC	671

Search completed: April 11, 2003, 23:49:11
Job time : 6920.21 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 03:06:43 ; Search time 7412 Seconds

(without alignments)
1849.355 Million cell updates/sec

Title: US-09-761-640-4

Perfect score: 2436
Sequence: 1 MALVTYSRSPGSGASTPVG.....PNEGFLROLIYOGILTART 471

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/cgnr_1/USPRO.spool/US09761640/runat_08042003.141434_20357/app-query.fasta.1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09761640.&CGN.1.14921.&runat.08042003.141434.20357 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEROJURY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-MARR_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_on: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pt: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_com: *
20: em_or: *
21: em_ov: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *

29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	2436	100.0	1416	6	AX451362 Sequence
2	2426	99.6	1980	6	AX451343 Sequence
3	2426	99.6	2781	6	AX099939 Sequence
4	2426	99.6	2808	9	AK074432 Sequence
5	2392	98.2	1416	9	AB072360 Homo sapi
6	2333	95.8	2905	9	AK094226 Homo sapi
7	1991.5	81.8	2736	10	BC028922 Mus muscu
8	1987.5	81.6	2322	6	AX398774 Sequence
9	1470.5	60.4	2064	6	AX405828 Sequence
10	1470.5	60.4	2064	9	AK000522 Sequence
11	1405.5	57.7	150903	2	AP002776 Homo sapi
12	1405.5	57.7	171566	2	AP001885 Homo sapi
13	1113	45.7	2280	9	AB072355 Homo sapi
14	1113	45.7	3817	9	AB072355 Homo sapi
15	1113	45.7	6374	6	AX180874 Sequence
16	1094.5	44.9	1796	9	AB072358 Homo sapi
17	1077	44.2	2433	9	AK095421 Homo sapi
18	1076.5	44.2	2260	6	AX180876 Sequence
19	1045	42.9	1711	6	AX223964 Sequence
20	1004.5	41.2	1771	6	AX202239 Sequence
21	994	40.8	1905	9	BC004176 Homo sapi
22	994	40.8	1905	9	BC004210 Homo sapi
23	980	40.2	1949	6	AX223960 Sequence
24	918.5	37.7	4417	3	AB036834 Drosophi1
25	852	35.0	3488	6	AX406972 Sequence
26	793	32.6	1026	6	AX086034 Sequence
27	760.5	31.2	1755	6	AX398768 Sequence
28	760.5	31.2	1755	9	AK001790 Homo sapi
29	733	30.1	181132	3	AC008206 Drosophi1
30	733	30.1	227398	3	AE003750 Drosophi1
31	699	28.7	113474	2	AC013932 Drosophi1
32	689	28.3	599	6	AX398772 Sequence
33	687	28.2	409	6	AX398771 Sequence
34	591.5	24.3	426	6	AX398765 Sequence
35	589	24.2	1052	6	AX223966 Sequence
36	426	17.5	5463	9	AB037719 Homo sapi
37	413	17.0	494	6	AX398773 Sequence
38	405	16.6	386	9	AF484838 Homo sapi
39	378	15.5	191571	2	AL845302 Dario rer
40	366.5	15.0	957	9	BC008941 Homo sapi
41	359	14.7	1008	9	AB072359 Homo sapi
42	322	13.2	935	9	BC011636 Homo sapi
43	310.5	12.7	1031	9	AB072357 Homo sapi
44	287	11.8	1083	8	AF116620 Arabidopsi
45	271	11.1	141017	2	AC116962 Dictyoste

ALIGNMENTS

RESULT 1

AX451362
LOCUS AX451362 1416 bp DNA Linear PAT 03-JUL-2002
DEFINITION Sequence 20 from Patent WO0224740.
ACCESSION AX451362
VERSION AX451362.1 GI:21698396
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dsp-15 dual-specificity phosphatase
TITLE Patent: WO 0224740-A 20 28-MAR-2002;
JOURNAL Ceptyr, Inc. (US)
FEATURES
source 1. 1416
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 288 a 450 c 449 g 229 t
ORIGIN
Alignment Scores:
Pred. No.: 3.88e-165 Length: 1416
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-761-640-4 (1-471) x AX451362 (1-1416)
QY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
DB 1 ATGGCCCTGGTACAGTGAAGCCGTTGCCCCGGGAGGGGCGCTCCAGCCGCGGGG 60
QY 21 ProTASPglnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
DB 61 CCTGGGACACAGCGGTTCCAGGAGAGAGTGCAGCTCCAGCGAAGGACAGCTTTGGCGT 120
QY 41 LeuAlaGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla 60
DB 121 CTCCTGGGGGCTGCTCGTGGAGTGCAGATGAGGGGACAAATGATCATGCACAGAGGCC 180
QY 61 SerSerGlnProThrGlnLysAlaProSerGlnGluGluLeuHisGlyAspGlnThrAsp 80
DB 181 AATTCTGAGCCACAGAGAGAGGCCCGAGTGAAGAGAGACTCCAGCGGACACAGACAG 240
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
DB 241 TTCGGGCAAGGATCCAGAGTCCCGCAGAGCAGAGAGAGGAGGACGACCTGCACCTC 300
QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGluAlaPro 120
DB 301 ATGGTACACACTCTGAGCGGAGGATGACATCCGCTGGCAGCCCGCAGCTGGAGGACCC 360
QY 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGlyLeuSer 140
DB 361 CGGCTCCCGGCGCTCCGCTGAGTTCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GlnAspGlnThrValLeuLeuGlnGlyValAspPheProAspSerSerSerProSerCysThr 160
DB 481 CTGGGCTGCTGCTGGCCCTGAGTGCACCCAGAGTACTTAAATGAGAGCGGGGCTC 540
QY 161 LeuGlyLeuValLeuProLeuTyrPheSerAspThrGlnValTyrLeuAspGlyAspGly 180
DB 181 PheSerValThrSerGlyGlnSerArgIlePhePhePheProIleSerIleGlnThrMet 200
QY 541 TTCAGGCTACGCTGCTGGTGGGCAAGCCGAGTCTTAAGCCCATCTCCATCCAGACATG 600
QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaLeuGlySerGlyLeuVal 220

DB 601 TGGCCACACTCCAGTATATGCACCAAGATGTGAGGACGCTCTAGGACGGGCTGTGA 660
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
DB 661 CCGGGTGGAGAGCCCTCAGCCAGCCAGCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnGluSerLeuArgProProSerAla 260
DB 721 AGCTGCTCAATGATGTGAGAGGCTATGGCCCGACCTGGAGATCTCTCGGCTCCAGCCGC 780
QY 261 GluProGlyGlySerSerGlnGlnGluGlnMetGlnGlnAlaIleArgAlaGluLeuTrp 280
DB 781 GAGCTGGCGGCTCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
DB 841 AAGGTGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
DB 901 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCTGACCTGATTCATGACACACAGATGCTG 960
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
DB 961 CTGCTGTGGGACACAGCGGAGCAGACCTCCGCAATCTCCACCTGATCTGAGCTGCTCA 1020
QY 341 GluTrpAsnAlaAlaAsnLeuGluGluLeuGlnArgAsnArgValThrHisTyrLeuAsn 360
DB 1021 GAGTGGAGACGACCAACCTGGAGAGGCTGCAGAGAGAACAGGGGTCAACCATCTTGAAC 1080
QY 361 MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
DB 1081 ATGGCCCGGAGATTTGACACTTACCTGAGCGCTTACCTACACAGATGGGCGCTC 1140
QY 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTyrPheGluThrHisArgPheIleGlu 400
DB 1141 TGGGATGAGAGTCCGCGCCAGCTGCTGCGCAGTGGAGAGAACCCAGCGCTTCAATTGAG 1200
QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
DB 1201 GCTGCAGAGACACAGAGGACCCACAGCTGCTGTGTCACATGCAGATGGGCTCAGCGCTCA 1260
QY 421 AlaAlaThrValLeuAlaValAlaMetLysGlnTyrGluCysSerLeuGlnAlaLeu 440
DB 1261 GCGGCACAGCTGCTGGCCATGCATGAAGCAGTACGAATGACAGCTGGAGCAGCGCTG 1320
QY 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
DB 1321 CGCCACGTGCAGAGACTCCGGCCCATTCGCCCGCCCAACCCCTGGGCTTCCTGGCGCAGCTG 1380
QY 461 GlnIleTyrGlnGlyTyrLeuThrAlaArgThr 471
DB 1381 CAGATCTACCAAGGACATCTGACGGCCAGAACCC 1413
RESULT 2
LOCUS AX451343 1980 bp DNA Linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0224740.
ACCESSION AX451343
VERSION AX451343.1 GI:21698394
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dsp-15 dual-specificity phosphatase
TITLE Patent: WO 0224740-A 1 28-MAR-2002;
JOURNAL Ceptyr, Inc. (US)
FEATURES
source 1. 1980
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 409 a 616 c 640 g 315 t
ORIGIN

Alignment Scores:
Pred. No.: 3.03e-164 Length: 1980
Score: 2426.00 Matches: 469
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.59% Indels: 0
DB: 6 Gaps: 0

US-09-761-640-4 (1-471) x AX451343 (1-1980)

```

Oy 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 1 ATGGCCCTGGTCAAGTACAGTACGCTTCCGCCCGGCGAGCGGCGCTCCACGCCGCTGGGG
Oy 21 ProTPASpGlnAlaValAlGlnArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Db 61 CCTTGAGACACAGCGGCTCCAGGAGAGAGTCACTCCAGCAGGACAGAGCTTTGGCGTG 120
Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla 60
Db 121 CTCCCTGGGGGCTGCTCGGAGCTCGAGATGGAGGGGCAATGATGATGACAGAGGCC 180
Oy 61 SerSerGlnProThrGlnLysAlaProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Db 181 AGTTCTGAGCCAAACAGAAAGAGGCCCCGAGTGAAGAGAGCTCCAGGGGACACAGAC 240
Oy 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
Db 241 TTCCGGGCAAGGATCCCAAGAGTCCCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Oy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaPro 120
Db 301 ATGTGACAGCTGCTGAGGCGCAGAGATGACATCCGCTGGACGAGCTGAGGAGGACACC 360
Oy 121 ArgProProAlaArgLeuArgTyrLeuValValSerThrArgGlnGlyGlnGlyLeuSer 140
Db 361 CGGCTCCCGGCGCTCCGCTACCTGCTGAGTCTTACACGAGAGAGAAAGGCTTGAGC 420
Oy 141 GlnAspGlnThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Db 421 CAGATGAGAGCGCTCTCTCGGCGTGATTTCCCTGACAGAGCTCCCCAGCTGCACAC 480
Oy 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
Db 481 CTGGGCGCTGCTTCCCTCGCTGAGTACACACAGAGTGTACTGATGAGAGAGGGGCGC 540
Oy 181 PheSerValThrSerGlyGlnSerArgIlePheIleProIleSerIleGlnThrMet 200
Db 541 TTCACCGGTGACTGTGGGGCAAGCCGAGTCTTCAAGCCCATCTCCATCCAGACAG 600
Oy 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
Db 601 TGGGACACACTCCAGTATGTCACCAAGATGTAGGCAAGCTCTAGGCGGCGCTTGA 660
Oy 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGlnAlaArgLeuAsnSerGln 240
Db 661 CCGGGTGGCAGTGCCTCACCCTGCGCAGCCACTACAGAGAGAGACTGAATCCGAGAC 720
Oy 241 SerCysLeuAsnGlnThrThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
Db 721 ACCTGCTCAATGAGTGGAGGCTATGGCCGACCTGGAGTCTCTCGGCTCCAGCGCC 780
Oy 261 GluProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 781 GAGCTGGGGGCTCTCAGAAACAGAGAGAGATGAGAGAGGAGATCCGCTCTAGCTGTGG 840
Oy 281 LysValLeuAspValSerAspLeuSerValThrSerLysGlnIleArgGlnAlaLeu 300

```

```

Db 841 AAAGTGTGGATGTCAGTACCTGGAGAGACTGTCACTCCAAAGAGATCCGCGAGCTCTG 900
Oy 301 GluLeuArgLeuGlyLeuProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
Db 901 GAGCTGGCGCTGGGCTCCCTCCAGCAAGTACCTGACTTATGACACACAGATGCTG 960
Oy 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
Db 961 CTGCTGGTGGCAGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Oy 341 GluTrpAsnAlaAlaAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 1021 GAGTGAAGCGCAAAACCTGGAGAGAGCTGCGAGAGAAACAGGCTCACCCATCTTGAAC 1080
Oy 361 MetAlaArgGlnIleAspAsnPheTyrProGluAspPheThrTyrHisAsnValArgLeu 380
Db 1081 ATGGCCCGGAGATTTGACACTTCTTACCTGAGCGCTTACCTACCAACATGTGGGCTTC 1140
Oy 381 TrpAspGlnGlnSerAlaGlnLeuLeuProHisTyrPlysGlnThrHisArgPheIleGln 400
Db 1141 TGGATGAGAGATTCGGCCAGCTGCTGCCGCACTGGAAGAGAGACAGCAGCTTCATGAG 1200
Oy 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1201 CCTGCAAGAGCACAGGCGCACCCAGCTGCTGTCACCTGCMAATGAGGCGCTCAGCCGCTCA 1260
Oy 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlnCysSerLeuGlnGlnAlaLeu 440
Db 1261 GCGGCCACAGTCTCGCTCCCTATCCATGAAGCAGTACGAAATGAGCTGGAGAGAGGCTCG 1320
Oy 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1321 CGCCACGTGACAGAGACTCGCGGCCATCGCGGCCGCCCAACCTGGCTTCTCGGCCACGTG 1380
Oy 461 GlnIleTyrGlnGlyIleLeuThrAla 469
Db 1381 CAGATCTACAGGCGCATCTCGAGGCGC 1407

RESULT 3
LOCUS AX099939 2781 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 21 from Patent WO0120004.
ACCESSION AX099939
VERSION AX099939.1 GI:13538949
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2781)
AUTHORS Yue, H., Tang, Y.T., Bandman, O., Hillman, J.L., Baughn, M.R.,
Azimzal, Y. and Lu, D.A.
Protein phosphatase and kinase proteins
Patent: WO 0120004-A 21 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
source 1..2781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID NO: 5039718CB1"
BASE COUNT 576 a 906 c 820 g 479 t
ORIGIN
Alignment Scores:
Pred. No.: 4.61e-164 Length: 2781
Score: 2426.00 Matches: 469
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.59% Indels: 0
DB: 6 Gaps: 0
US-09-761-640-4 (1-471) x AX099939 (1-2781)

```

QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyValAsnThrProValGly 20
 DB 87 ATGGCCCTGGTCACTGAGCCCTTGGCCCCGGGAGCGGCGCTCCACAGCCCGTGGGG 146
 QY 21 ProTPrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
 DB 147 CCTGGGACAGCGGTGTCAGCCAGAGAGTCTCCAGGAGGAGGAGACCTTTGGGTG 206
 QY 41 LeuArgGlyValValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
 DB 207 CTCCTGGGGGCTCTCTGGGACCTGACAGATGAGGGGACATATATATGACAGAGGCC 266
 QY 61 SerSerGluProThrGluLysAlaProSerGluGluGluGlnAspGlnThrAsp 80
 DB 267 AGTTGAGACCAACAGAGAAGGCCCGGAGTGAAGAGAGAGCTCCACGGGGACAGACAG 326
 QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeu 100
 DB 327 TTCGGGCAAGAGATCCAGAGTCCCGAAGCAGAGAGAGAGGAGGACCTGACCTC 386
 QY 101 MetValGlnLeuLeuArgProGlnAspAlaArgLeuAlaAlaGlnLeuGlnAlaPro 120
 DB 387 ATGCTACACTGCTGAGCGCGGACGATGACATCCGCTGGACCCAGCTGGAGGACCC 446
 QY 121 ArgProProArgLeuArgGlyLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSer 140
 DB 447 CGGCTCCCGGCTCGCTACCTGCTGATGTTCTACAGAGAGAGAGAGAGGCTGAGC 506
 QY 141 GlnAspGlnThrValLeuGlyValAspPheProAspSerSerSerProSerGlyThr 160
 DB 507 CAGAGTGAAGAGGCTCTCTCGGGCGTGAATTCCTCAGACGCGTCCCGACCTCCACC 566
 QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValThrLeuAspGlyAspGly 180
 DB 567 CTGGGCTGCTGTCGCCCTCTGAGTGAACCCAGCTGACTTACATGAGAGCGGGGC 626
 QY 181 PheSerValThrSerGlyGlnSerArgGlyPheLeuSerProLeuSerLeuGlnThrMet 200
 DB 627 TTCAGGCTACGCTGCTGGGCAACCCGAGTCTTCAAGCCCTTCATCCAGACCATG 686
 QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
 DB 687 TGGGCCACACTCCAGATATTGCACCAAGATGTGAGGCGCTTAGCGAGCGCTTGTG 746
 QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGlnArgLeuAsnSerGlnGln 240
 DB 747 CCGGGTGGAGTCCCTCACTGAGCCAGCCACTACAGAGAGACTGAACCTCCGAACG 806
 QY 241 SerCysLeuAsnGlnThrPheAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
 DB 807 AACTGCTCAATGAGTGAAGCGGCTATGGCCGACCTGAGTCTCTCGGCTCCCACTG 866
 QY 261 GluProGlyGlySerSerGlnGlnGlnGlnMetGlnGlnAlaLeuArgAlaGlnLeuTrp 280
 DB 867 GAGCTGGGGGCTCTCAACAACGAGGAGATGAGACAGCGATCCGCTGAGCTGG 926
 QY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerLysGlnLeuArgGlnAlaLeu 300
 DB 927 AAAGTGTGAGTGCAGTACCTGGAGACTGTCACTTCCAAGAGATCCGCCAGGCTG 986
 QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheLeuAsnGlnMetLeu 320
 DB 987 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGTACCTCATGCAACACGATGCTG 1046
 QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgGlyLeuPheProHisLeuThrLeuGlySer 340
 DB 1047 CTGCTGGTGGCAAGGGGAGCGAGCTCCCGATCTTCCCACTTACCTGGGCTCA 1106
 QY 341 GluTrpAsnAlaAlaAsnGlnGlnLeuGlnArgAspArgValThrHisLeuAsn 360
 DB 1107 GAGTGAAGCGACCAACTGCGAGGCTGCGAGAGAAAGGCTCCACATCTTTGAAC 1166
 QY 361 MetAlaArgGlnLeuAspAsnPheTyrProGluArgPheThrHisAsnValArgLeu 380

DB 1167 ATGGCCGGGAGATTGACAACCTTCACTGAGCGGCTTCACTACCAACAATGAGCGCTTC 1226
 QY 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTyrLysGlnThrHisArgPheLeuGlu 400
 DB 1227 TGGGATGAGAGTGGCGCCAGCTGCTGCCGACTGGAGAGAGCGGACCGCTTCATTTGAG 1286
 QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
 DB 1287 GCTGCAAGAGCAGAGGCGACCCAGCTGCTGCTCCTGCTCAAGATGAGGCGCTGAGCCCTTA 1346
 QY 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnAlaLeu 440
 DB 1347 GCGCCACAGTCTGCTGCTATGCTATGAGAGATGAGATGAGCTGAGAGCAGGCGCTG 1406
 QY 441 ArgHisValGlnGlnLeuArgProLeuAlaArgProAspProGlyPheLeuArgGlnLeu 460
 DB 1407 CGCCAGTGCAGAGAGCTCGGCGCCATCGCCGCCACCCCTGCTTCTGGCGCAGCTG 1466
 QY 461 GlnIleTyrGlnGlyIleLeuThrAla 469
 DB 1467 CAGATTCACAGGCGCATCTGACGCGC 1493
 RESULT 4
 AK074432
 LOCUS
 DEFINITION
 AK074432
 AK074432.1 GI:18677033
 VERSION
 AK074432.1 GI:18677033
 KEYWORDS
 Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone_11b:KAT clone:KAT12021.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1
 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 TITLE
 JOURNAL
 JOURNAL
 AUTHORS
 2 (bases 1 to 2808)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
 (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction: 5'-& 3'-end one pass sequencing: Department of Vitrology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
 location/Qualifiers
 1..2808
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="KAT12021"
 /cell_line="KATO III"
 /cell_type="signet-ring cell carcinoma"
 /clone_lib="KAT"
 /note="cloning vector: pME18SFL3"
 97..2076
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB85080.1"
 /db_xref="GI:18677034"
 /translation="MALYTVRSPPGCSASIPVGPMDQAVQRSRRLQRQSFAYLRGA VLGLODGDNDDAEAFASSEPFERKAPSEFEELGDDTDFQSGSOSPOKDEQKHLLWV QLRPODQIRLAQOLEARPPRLRLVSVTRREGSLQDETIVLLGVDFPSSPSCV

LGIVLPLMSDPOVYVLDGDGGSVTSNGASRFKRFKISITQTMATLTOHCEALHSG
 LVPGGSLVTWASHYOERLUNBSOCLNENGTATADLESLEBPSCAFEGGSEBCEOMQALH
 AELMKVLDVDSLESVTSKEIRLOALELRIGTLPLOOYRFPIDNOMLLVAQRASRIFFP
 HLYTSEWMNANLEELORNYTHILNNAREDINYPEPFTYHNYLMDSEASQLPHK
 KETHRFTEAARAOGTIVLHCKMGVSSAAVLYAMQVCSLEALRHOVELRPIR
 RPNNGFLBOLDITOGILTLASRQSHVMECKQGVSPDEHPAPEVSTPEPFLPEPDEGGG
 EEKVLVMEESQAPAKEEPGRPRINLRPMKSSILLESVTSLESTSESDMPWEVSSIS
 EESSHEEPLQEPOLARTKGGQOVDRCGRPOVMSISLVSSTLSSAVNARTQAFQEBE
 GGQGGGQEPCCSSNTPRSRKKVYRQASVHDSGEGEA"

Pred. No.:	4,666-164	Length:	2808
Score:	2426.00	Matches:	469
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	99.59%	Indels:	0
DB:	9	Gaps:	0

US-09-761-640-4 (1-471) x AK074432 (1-2808)

Oy	1	MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly	20
Db	97	ATGGCCCTGGGTACAGTAGGCCGTCGCCGCCGGGCGAGCGGCGCTCCACGGCCGGGGG	156
Oy	21	ProTPASPGLInAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal	40
Db	157	CCCTGGGACACAGCGCGGTCCAGCGAAGAGATGCACCTCCACGAAGCAGAGCTTTGGGGT	216
Oy	41	LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla	60
Db	217	CTCCCTGGGGCGCTGCTCGGAGCTGCAGCATGGAGGGGACAAATGATGTCACACAGAGCC	276
Oy	61	SerSerGlnProThrGlnLysAlaProSerGlnGlnGlnLeuLeuHisGlyLysProGlnThrAsp	80
Db	277	ACTTCTGAGCCCAACAGAGAGGCCCGCGAGTGAAGAGAGACTCCACGGGGACCAAGACAGAC	336
Oy	81	PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnArgGlnHisLeu	100
Db	337	TTCCGGGCAAGAGATCCCAAGATGCCCAAGACAGAGAGACAGAGGAGAGACCTGCACCTTC	396
Oy	101	MetValGlnLeuLeuArgProGlnAspAspLLeuArgLeuAlaAlaGlnLeuGluAlaPro	120
Db	397	ATGGTAACACTCTGAGGCGCGAGATGCATCCGCTGGACGCCAGCTGGAGGACACCC	456
Oy	121	ArgProProArgLeuArgArgArgLeuValValSerThrArgGlnGlyLysGlyLeuSer	140
Db	457	CGGCGTCCCGGCTCCGCTACCTGGCTGGATTTCTTACACGGAAGAGAAAGATCTGGAGC	516
Oy	141	GlnAspGlnThrValLeuLeuGlnLysAlaAspPheProAspSerSerSerProSerCysThr	160
Db	517	CAGGATGAGACGGTCTCTCGGGCGTGGATTTCCCTGACAGCAGCTCCCCACGCTCCACC	576
Oy	161	LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly	180
Db	577	CTGGGCGTGGTCTGGCCCTCTGAGATGCACCCACAGGTACTAGATGAGACGGGGGCG	636
Oy	181	PheSerValThrSerGlyLysGlnSerArgLLeuPheLysProLLeuSerLLeuGlnThrMet	200
Db	637	TTTCAGGGTACCTCTGGTGGGCAAAACCGGATCTTCAACCCCATCTCCATCCAGACCAG	696
Oy	201	TyrAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal	220
Db	697	TGGGCCACACTCCAGGTATTGCACCAAGATGTGAGGACGCTTAGGACACGGCCTTGTGA	756
Oy	221	ProGlyGlySerAlaLeuThrTrrAlaSerHisTyrGlnGlnArgLeuAsnSerGlnGln	240
Db	757	CCGGGGCAGTGGCCCTCACTCGGGCCACCCTACCAAGAGAGACTGAACCTCCGAACAG	816
Oy	241	SerCysLeuAsnGlnThrPheThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla	260
Db	817	ACCTGCTCTAAATGAGTGAGCGGCTATGAGCCGACCTGGAAGTCTCTTGGGGCTCCACAGCC	876

QY	261	GIUProGlyGlySerSerGIUGlnGlnMetGlnGlnAlaIleArgAlaGlnLeu	280
Db	877	GAGCCTGGCGGGCTCTCCAGAACAGGAGAGATGGAGCGCATCCGGCTGAGCTGG	936
QY	281	LysValIleAspValSerAspLeuGlnSerValThrSerIysGluIleArgGlnAlaLeu	300
Db	937	AAAGCTTGAGATGTCAGTACGACCTGGAGAGTGCATCTCCAAAGAGATCCGGCAGGCTTG	996
QY	301	GIULeuArgLeuGlyLeuProLeuGlnGlnIleThrArgAspPheIleAspGlnMetLeu	320
Db	997	GAGCTGCCTCGGGGCTCCCTCCAGCAGTACCTGACTTCATCGAACACAGATGCTG	1056
QY	321	LeuLeuValAlaGlnArgAspArgLaseArgIlePheProHisLeuThrLeuGlySer	340
Db	1057	CTGGCTGGGGCCACAGCGGAGCCGAGCCTCCCGCATCTTCCCCCACTCTTACTGGGCTCA	1116
QY	341	GluTPAsnAlaIleAsnLeuGlnGlnLeuGlnArgAsnArgValThrHisIleLeuAsn	360
Db	1117	GAGTGAACGCGCAGCAACCTGGAGAGCTGCAGAGAACAGGTCACCCACATCTTGAAC	1176
QY	361	MetAlaIleArgGluIleAspAsnPheArgProGluArgPheThrThrHisAsnValArgLeu	380
Db	1177	ATGGCCCGGGAGATTGACAACTTACCTCGAGCCTTCACCTACCACAAATGTGGCCTC	1236
QY	381	TRPAspGluGlnSerAlaGlnLeuLeuProHisPheIysGluThrHisArgPheIleGlu	400
Db	1237	TGGGATGAGAGATCGGGCCAGCTGCTGCCGCCACTGGAGAGAGACCCAGCGCTTACTAG	1296
QY	401	AlaAlaIleArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer	420
Db	1297	GCTGCAGAGGACACAGGAGACCCAGCCTGTGGTCCACTGCAGATGGGGTCCAGCGCTCA	1356
QY	421	AlaAlaIleThrValLeuAlaIleThrAlaMetIysGlnIleArgIleCysSerLeuGlnAlaLeu	440
Db	1357	GCGGCCACAGTCTGGCCTATGCTCATGAGACGATGCGANTGCAAGCTGGAGCAGGCCCTG	1416
QY	441	ArgHisValGlnGlnLeuLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu	460
Db	1417	CGCCACGTCAGAGAGCTCGGCCCATCGGCCGCCCAACCTGGGCTCTCGTGGCAGCTG	1476
QY	461	GlnIleTyrGlnGlyIleLeuThrAla 469	
Db	1477	CAGATCTACAGGAGCATCTGAGGCC 1503	

RESULT 5

LOCUS	AB072360	1416 bp	linear	PRI 29-JAN-2002
DEFINITION	Homo sapiens mRNA for hssh-3, complete cds.			

VERSION	AB072360.1	GI:18376668
KEYWORDS		
SOURCE	Homo sapiens	CDNA to mRNA

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

MEDLINE	21822082
REFERENCE	2 (bases 1 to 1416)
AUTHORS	Niwa, R., Nagata-Ohashi, K., Hay, B.A., Takeichi, M., Mizuno, K. and

FIELD DIRECT SUBMISSIONS
Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular Genetics, The Institute for Virus Research, Kyoto University; Shogoin-Tawara-cho 53, Sakyo-ku, Kyoto 606-8507, Japan
(E-mail: tucumura@virus.kyoto-u.ac.jp, Tel: 81-75-751-4031, Fax: 81-75-751-3989)

FEATURES	Location/Qualifiers
source	1. .1416

gene
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1416
/gene="hSSH-3"
1..1416
/gene="hSSH-3"
/standard_name="slingshot-3"
/note="a putative dual specific phosphatase"
/codon_start=1
/product="hSSH-3"
/protein_id="BAB84119.1"
/db_xref="GI:18376669"
/translation="MALVYSGSPGSGASTPVGPMDOAVQRRSLQRQSFAYLGA
VLGDDGNDNDVAVSESPTEKAPSEELHGDQDFQSGSQSPQKEOROHLLMY
OLRPFODIRLAQOLEARPPRIYLVSTREGELSGDETVLLGDFPDSSPSC
LGLVPLMSDVOYLDGDFGFSVTSGGSSRTFKPISIQTMATLOVLRQACEALGSG
LVPGSALTWASHYQERLNSQSLNEMTMADLESRLPRSPRSGSEODMEQAIR
AEIMKVLVSDLESVTSKEIRQJWSCAMGSPSSSTVISTTRCCCMHMSQDRASIRP
HLVIGSEWNAANLEELQNRVYHILNMAREDINFPERTYHNHVLNMEESQQLPIA
KTHRFTEARAGQTHVLVHCHCMGVSRSRAVLAYAMKQYECSEQALRHVQELRPIA
RNPSPFLRQLOIYQGLTART"

BASE COUNT 288 a 449 c 449 g 230 t
ORIGIN

Alignment Scores:
Pred. No.: 5, 39e-162 Length: 1416
Score: 2392.00 Matches: 469
Percent Similarity: 99.36% Conservative: 0
Best Local Similarity: 99.36% Mismatches: 2
Query Match: 98.19% Indels: 0
Gaps: 0

US-09-761-640-4 (1-471) x AB072360 (1-1416)

OY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 1 ATGCCCTGGTGCACAGTGAAGCGCTTCCGCCCGGCGACGGCGCTCCAGCCCGTGGGG 60
OY 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
Db 61 CCGTGGGACAGGGGGGTCCAGCAGAGCTCCAGCCAGGCAAGCAAGCTTGGGGTGG 120
OY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyAspAsnAspAlaAlaGlnAla 60
Db 121 CTCCCTGGGGGTCTCCCTGAGCTGAGATGAGAGGGCAATGATGATGAGATGAGAGCC 180
OY 61 SerSerGluProThrGlnLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
Db 181 AGTTCTGAGCCAAAGAGAGAGCCCGAGTGAAGAGAGCTCCAGCGGAGCCAGCAGAC 240
OY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
Db 241 TTCGGGCAAGGATCCAGAGTCCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 101 MetValGlnLeuLeuArgProGlnAspAspLeuArgLeuAlaAlaGlnLeuAlaPro 120
Db 301 ATGGTACACCTCTCTGAGCGCGAGATGATCCGCTGCGACGCCAGCTGAGAGGACACC 360
OY 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGluGluLeuSer 140
Db 361 CGGCGTCCCGGCTCCGCTACCTGCTGATGTTCTTACACGAGAAAGAGAGAGTCTGAGC 420
OY 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Db 421 CAGAGTGAAGACGGTCTCTCCGGGGCTGGGATTTCTCTGACAGCGAGCTCCGCCACTGACC 480
OY 161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
Db 481 CTGGGGCTGCTCTTGGCCCTCTGAGTGAACCCAGAGTGAATAGTGAAGAGCGGGGCGC 540
OY 181 PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
Db 541 TTCAGCGTGAAGTCTGGTGGGCAAGCGGATCTTCAAGCCATCTCCATCCAGACCATTG 600

OY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
Db 601 TGGCGCACACTCCAGATGATTGTCACCAAGCAGTGAAGGAGCTCTTGAAGCGGCTTGTGA 660
OY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
Db 661 CCGGGTGGCAGAGCCCTTCACTGGGCCAGCCACATACAGAGAGACTGTAACCTCCAGACAG 720
OY 241 SerCysLeuAsnGluTyrTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
Db 721 AGCTCCCTCAATGATGATGAGCAGCGGCTATGGCCACCTGGAGTCTTGGGGCTCCAGAGCC 780
OY 261 GluProGlyGlySerSerGluGlnGlnGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
Db 781 GAGCTGGCGGGTCTCTCAGACAGAGAGCAGATGAGCAGAGCGATCCGCTGAGCTGTGG 840
OY 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
Db 841 AAGTGTGTGATGTCAGTCACTGAGAGATGTTCACCTTCCAAAGAGATCCGCCAGCT-CTG 899
OY 301 GluLeuArgLeuGlyLeuLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
Db 900 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCTGACTTCATGACAGACCAAGATGCTG 959
OY 321 LeuLeuValAlaGlnArg-AspArgAlaSerArgIlePheProHisLeuTyrLeuGlyLys 340
Db 960 CTGCTGGTGGCAGAGCGGGGAGCGAGCTCCCGCATCTTCCCGACCTTACTTGGGCTTC 1019
OY 340 rGluTrpAsnAlaAlaAsnLeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAs 360
Db 1020 AGAGTGAAGCAGCAAACTGGAGAGGCTGCGAGAGAGACAGGGTCAACCATCTTGA 1079
OY 360 mMetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLe 380
Db 1080 CATGGCCCGGAGATTCACAACTTACCTGAGCGCTTACCTACACAACTGGCGCT 1139
OY 380 uTrpAspGluGluSerAlaGlnLeuLeuProHisTyrLysGluThrHisArgPheIleG 400
Db 1140 CTGGGATGAGAGTCCGCCAGCTGCTGCGGCAGCTGGAAGAGAGCAGCGCTTCAATGA 1199
OY 400 uAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSe 420
Db 1200 GGGTGCAGAGACACAGGACACCCAGCTGCTGCTCCTGCAAGATGGGCTCAGCCCTC 1259
OY 420 rAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnAlaLe 440
Db 1260 AGCGGCCACAGTCTGGCTGTGCTATGCCATGAACAGTACGAATGACAGCTGGAGCGCCCT 1319
OY 440 uArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLe 460
Db 1320 GCGCCAGCTGAAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGGCGCAGCT 1379
OY 460 uGlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
Db 1380 GCAGATCTACCAAGGCAATCTGACGCGCAGAAC 1413

RESULT 6
AK094226 2905 bp mRNA linear PRI 15-JUN-2002
LOCUS Homo sapiens cDNA FJ36907 f1s, clone BRACE2003800, weakly similar
DEFINITION to MAP kinase phosphatase.

ACCESSION AK094226
VERSION AK094226.1 GI:21753246
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_11b:BRACE2
clone:BRACE2003800.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiroo,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
AUTHORS

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H.-I., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, R., Nakamura, Y., Sekine, M., Kituchin, H., Kanda, K., Wagatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2905)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUN-2002) Takao Isogai, Fuj. Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomeshcrt.co.jp, Tel: 01-438-52-3975, Fax: 01-438-52-3986)
 NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source

CDS

BASE COUNT	591 a	940 c	867 g	507 t
ORIGIN				

Alignment Scores:

Pred. No.:	2,13e+157	Length:	2905
Score:	2333.00	Matches:	467
Percent Similarity:	89.98%	Conservative:	0
Best Local Similarity:	89.98%	Mismatches:	1
Query Match:	95.77%	Indels:	32
DB:	9	Gaps:	3

US-09-761-640-4 (1-471) x AK094226 (1-2905)

OY	1	MetAlaLeuValThrValSerArgSerProProIlyserGlyAlaSerThrProValGly	20
Db	64	ATGGCCCTGGTTCACAGTGGCCGTTCCGCCCCGGGACAGGGCGCTCCACGGCCCGTGGG	123
OY	21	ProTrp-----	22
Db	124	CCCTGGAACTTCCTGAGAGAGGGGAGACAGCCCTCCCGCCCTCACCTGCCCTTGGG	183
OY	23	-----AspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSer	37
Db	184	CTGCTCTCTCGGACAGACAGCGGCTCCAGCGAAGAGATGACTCCACGAGAAAGT----	239
OY	38	PheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAla	57
Db	240	TTTGGGCGTCTCGTGGGGCTGTCTCTGGACATGCAGATGAGAGGGGCAATGATGATGCA	299

QY	58	AlaGluAlaSerSerGluProThnGluValaProSerGluGluGluLeuHisIseLysP	77
Db	300	GCAGAGGGCACTTGTGAGCCAAACAGAAAGAGCCCCAGAGAGAGAGCTCCACGGGAGC	359
QY	78	GlnThrAspPheIysGlnIysSerGlnSerProGlnIysGlnGluGlnIArgIHis	97
Db	360	CAGACAGACTTCGGGCAAGAGATCCAGATCCCCACAAACAGAGAGAGCAGAGGCAGAC	419
QY	98	LeuHisLeuMetValGlnLeuLeuValProGlnAspAspIleArgLeuAlaIaGlnLeu	117
Db	420	CTGACCCCTCAGGTACAGCTGCTGAGGCCAGAGATCCGCTGGCACCCAGCTG	479
QY	118	GluAlaProArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGlu	137
Db	480	GAGGACCCCGCGCTCCCGGCTCCGTACTCTCTGTAGTTCTACACAGAAAGAGAGA	539
QY	138	GlyLeuSerGlnAspGluThrValLeuLeuGlyValAspPheProAspSer-----	154
Db	540	GGTCTGACCCAGAGATGAGAGGTCCTCTCGGCGTGGATTTCCCTGCACAGAGTTGCAG	599
QY	154	-----	154

Db 600 CAGGGAAGGAAGGAGGGGCAAGAGAGGTGAGGCAACAGGGGGGCAATGGCTTTCCCTCT 659

QY 391 strplysgluthrhisarphellleclualalaaraglainglythrhisvalleu 411
|||||
Db 1380 CTGGAAGGAGACACCGCTTCATTCAGCTGCACAGAGCAGCAGCAGCAGCTGCTGT 1439
QY 411 1HsCysLysMetGlyValSerSerSerAlaAlaThrValLeuAlaTyrAlaMetLysG 431
|||||
Db 1440 CCAGCTCAAGATGGGGCTACGGCTCAGCGGCGCACAGTCTGCTATGCCATGAAGCA 1499
QY 431 nryglucysserleugluinalaleuarqhisvalaglingleuleuarprollealaar 451
|||||
Db 1500 GTACGATGACAGCTGAGAGAGGCGCTCGCCACAGTCGACAGAGACTCCGGCCATCGCCG 1559
QY 451 gproasnpoglypheleuarqingleuqinletyryglingleuleuthala 469
|||||
Db 1560 CCCCAACCTGGCTCTCGCGCAGCTGCAGATCTACAGAGGATCTCTGACGCGC 1614
RESULT 7
LOCUS BC028922 2736 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, similar to slingshot 3, clone MGC:25738
IMAGE:3987714, mRNA, complete cds.
ACCESSION BC028922
VERSION BC028922.1 GI:20810382
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2736)
REFERENCE 1
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapds-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarthe, P.H., Garcia, A.M., Lu, X., Huliyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 30 Row: 9 Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 2736
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZCH IT"
/clone="MGC:25738 IMAGE:3987714"
/tissue_type="Mammary tumor metastasized to lung, tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMTV."
/clone_lib="NCT_CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
93..2042
/codon_start=1
/product="similar to slingshot 3"
/protein_id="AAH28922.1"
/db_xref="GI:20810383"

/translation="MALVTSKSPASGHPVEGPTODRVVRRGRIRORSAFVLRG
AVLGLODGDGNVASEADSEPMEEPSGEDEPTEDTDGGLQSPDWKQVRHLAW
ELLRPDDIRLAOLEAPPLRLYLIVSGEELPEEPAIIIGDPSSSHCTGL
VLPMSPDVOYVIGDGGFSVTSGGOSIRFKPVSTIOTMTATLOVLIHOACVLAGGIVP
GGSALAMTYYVOEKINSQOGLNEMAMASLDSEFPRAEPFGQASBOEKMSQITLAE
MOVDLTSDSDVTSKEIRQALELRIGCPLOQTRDFTIDMQLLLMAQDDRASRITPPL
LGSEWMANLELELOKNRVSHILNMARELIDNPEPFTYNNRVVDEESAOLIPHKET
HRFTEDARAOGTRVYLHCKMGVSRAAIVLAVAMKQXGMDLEOALIHVOELPIVRPN
HGFLEOLRTYOGILITASROSHVMEQKVVSPEEPLAEVSTPLPEPFGSGGEVM
VMGLEDSOETPEKEELGLRPIRLNRLVMSISILEPSESTPEAGGLPEVSPDEPL
HPFOLSRAKGQQRVRRKQPMWALKRSQSVVALHSAVALYASNTRAFQBOGQOBOSEFG
MSSTPRLRKRMKQASVDSREDDKA"
BASE COUNT 599 a 780 c 768 g 589 t
ORIGIN
Alignment Scores:
Pred. No.: 4 84e-133 Length: 2736
Score: 1991.50 Matches: 394
Percent Similarity: 90.43% Conservative: 31
Best Local Similarity: 83.83% Mismatches: 40
Query Match: 81.75% Indels: 5
DB: 10 Gaps: 4
US-09-761-640-4 (1-471) x BC028922 (1-2736)
QY 1 MetaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
|||||
Db 93 ATGCGCTGTCACAGTGAATGCTGCGCGCGCAAGTGGCCACATCCAGCTGGGGA 152
QY 21 Pro---TriPspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAla 39
|||||
Db 153 CCCAGCAGGACCGAAGTGTAAAGCAGACAGCGGCTCCACGCGAGGAGGATTTGCG 212
QY 40 ValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaGln 59
|||||
Db 213 GTCTCTCCAGAGGCGCTCTCTGAGACTTCAGATGAGAGGAGAGATGATTCGACTGAG 272
QY 60 AlaSerSerGluProThrGluLysAlaProSerGluGluLeuHisGlyAspGlnThr 79
|||||
Db 273 GCTGACTGTGAGCAATGAGAGAACCTCAGGTGAGAGACCCCGAGGAGCAAGACT 332
QY 80 AspPheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHis 99
|||||
Db 333 GACAAAGGCGCAAGACTCAGAGTCCCTGGAACAA---GTCCAAAGAGGAGCCTGCGAC 389
QY 100 LeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAla 119
|||||
Db 390 CTCATGTGTGAGCTACTGAGCGCGCAGATGATATCCGCTGGACGCCAGCTGAGAGCA 449
QY 120 ProArgProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGlyLeu 139
|||||
Db 450 GCCCGACCCCGCAGACTGCTGATCGTGTGATGATTCACG-----GGAGAAAGACTG 503
QY 140 SerGlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerProSerCys 159
|||||
Db 504 AGT---GAGAGGCGCATCTGCTGGGGTGACTTCCCGAAGCAGTTCCCAAGAGCTTC 560
QY 160 ThrLeuGlyLeuValLeuProLeuThrPheSerAspThrGlnValTyrLeuAspGlyAsp 179
|||||
Db 561 ACCCTGGCGCTGCTTGTGCTTGTGAGTATACCAAGGTACTTACCTGAGCGGATGGG 620
QY 180 GlyPheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThr 199
|||||
Db 621 GGCTTAGAGTGACATCTGTGTCAAAACCGAATCTTCAACCGTCTCTATCCAGAC 680
QY 200 MetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeu 219
|||||
Db 681 ATCTGGCCCACTGCTGATGACCAAGCATGAGATGAGATGAGTCTAGCGAGTGGCTTT 740
QY 220 ValProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluAlaGlyLeuSerGlu 239
|||||
Db 741 GTGCTGTGGGAGTGCTCCCTGCTGCTGCTACTTACCAAGAGAACTCAACTCTGAC 800

OY	240	GINSERCYSLEASNSGUITRPHRLAMETALASPRELUGUSELEUARPROPOSER	259		
Db	801	CAGGGTTCCTCAAGAATGTGAATGGCCATTGCTTGACCGTAGGATCCTTTGCCACACTTAAT	860		
OY	260	AAGAUPROGLIYLSERSESGIUNGLINMETGLUGINALIAARGAAGLUDEU	279		
Db	861	GCTGAGCCTGGACACAGGCTTCCAGACAGAAAAGATGAGCACAGGCATCTTGCTGAAATTG	920		
OY	280	TRPLYSVALLEUSAPVALSERASPBLEUGUSERVALJTHRSERYSGUILEARGLNALA	299		
Db	921	TGGCAGAGGTGTGGACACCAGTGAAGCTGACAGACGGTCACCTTCAAGAGATCCGTGAGGCC	980		
OY	300	LEUGLULEUAIGLEUGLYLEUPROLEUNGLINTRYATARGSPHEITLASPANSIMET	319		
Db	981	CTGGAGCGCTCTGGGATGTCTCTCCAGCAGTACCTGACTTATATGACAACACAGATG	1040		
OY	320	LEULEUVALAAGHARARGAPRGALASERTGILPERNONISLEUTYTREUGLY	339		
Db	1041	TTGCGCTCATATGGCCCCAACAAACAGCCGGCCCTCTCCATCTTCCCACCTCTACTTGGGC	1100		
OY	340	SERGUTRPASNALAAALASNLEUGLULEUGLINARVASNARGVALTHISILEUEU	359		
Db	1101	TCTGATGTGAATCGTCCCACTGGAGGAACTTCAGAAAACAGATAGTACACTTCTG	1160		
OY	360	ASNMCETLAARGGLULIASPSANDHEYTPROGUATGRPHERHTYRHISANVALARG	379		
Db	1161	AACATGCGCCCGGAGATGTACAACTTCTCCCTGAGCGCTTACCTATTAACAATAGCT	1220		
OY	380	LEUTRAPSPGLUGUSERALAGLNULEUPRONISTPRLYSGLUTHNHISARGHELLE	399		
Db	1221	GCTGTGGATGAGAAATCGGCACAGCTTGTGCCCTGGAAAGAACACATCGATTCAAT	1280		
OY	400	GLUALAALARGAAGLAGLGLYTHRHSIVALLEUVALHISCYLSYMEGUYALISERTG	419		
Db	1281	GAGGATGCCAAGACGACAGAGGCACTGGGGTGTGATGCCACTGTMAAATGGGTCTGAGCGT	1340		
OY	420	SERIALAATHRVALLLEUNALETALAMELYSGINTRYGLUCYSERLEUGLUINALA	439		
Db	1341	TCTGCCCGCCACAGGCTACTGAGCCATGCTCAATAACAGTATGGCTGGAGCAACGCC	1400		
OY	440	LEUATGHLSVAILNGLULEUARPROILLEALAARGPROASNPROGLYPHELEUNARGGIN	459		
Db	1401	CTGATCCACGTGCAGGAGACTCGGCCCATGTGACGCCCCAACACACGAGCTTCTGCGCCAG	1460		
OY	460	LEUGINLIETYRGINGLYILEUTHRALA	469		
Db	1461	CTACGCGACTTACACAGGCATTTCTGACTGCC	1490		
RESULT 8	AX398774	2322 bp	DNA	linear	PAT 27-MAY-2002
LOCUS	AX398774				
DEFINITION	Sequence 10 from Patent WO0220732.				
ACCESSION	AX398774				
VERSION	AX398774.1	GI:21261307			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Liou,J.R.				
AUTHORS	1				
JOURNAL	Regulation of human map kinase phosphatase-like enzyme Patent: WO 0220732-A 10 14-MAR-2002; Bayer Aktiengesellschaft (DE)				
FEATURES	Location/Qualifiers				
	1..2322				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	467 a	751 c	718 g	386 t	
ORIGIN					
Alignment Scores:	7.64e-133		Length:	2322	
Pred. No.:					

Score:	1987.50	Matches:	428
Percent Similarity:	86.63%	Conservative:	6
Best local Similarity:	85.43%	Mismatches:	27
Query Match:	81.59%	Indels:	42
DB:	6	Gaps:	8

US-09-761-640-4 (1-471) x AX398774 (1-2322)

QY	4	ValThrValSerArgSerProPro-----	11
DB	53	GTCACAGTAGAGCGCTGCGCCCGCGGAGAGCTGGCGGCTCCACCGCCCTGTGGGGCCCTG	112
QY	12	-----GlySerGlyAlaSerThrProValGlyProThrPasp-----	23
DB	113	GAATTCCCTGAGAGAGGGAGGGAGACGCCCTCCCGCCCTCAGCCCTGGGG-CTGCTCTCT	172
QY	24	-----GlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPhaAlaValLeu	41
DB	172	CGGCAGGACCTGGCTCCAGCAAGAGGTGTGACTCCAGCCAGAA------GCTTTG	215
QY	42	Arg-----GlyAlaValLeuGlyLeuGlnGlnAspGlyGlyAspAsnAspAlaAla	58
DB	220	CGGTCTCTCTGTGGGGGCTGTCTCTGGGACTGCAGAGTGGAGGGACATGATGATCCAGCA	279
QY	59	GluAlaSerSerGluProThrGlnLysAlaProSerGlnGlnGlnLeuHis--GlyAsp	77
DB	280	GAGGCCAGTGTTCAGCCAAACAGAAAGCCCCCGATGGAGAGAGTCCACGGGGGAACC	339
QY	78	GlnThrAsp-PheGly-GlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnH	97
DB	340	CAGACAGACTTTCGGTGCAGAAAGATCCCGAGTCCCCGAAACAGAGAGAGAGAGGAGC	399
QY	97	IleLeuHisLeuMetValGlnLeuLeuArgProGln-AspAspIleArgLeuAlaAla---	115
DB	400	AACTCCAAAC--TCATGAGAGCGCTCTGAGGCGCCGACGGATGCATCCGACTTGGAAAGCCCA	458
QY	116	GlnLeuGlnAlaProArgProProArg-LeuArgTyr---LeuLeuValAlaSerThrAr	134
DB	459	AGCTGAGAGCAACCCCGGGGCTCCCGGGATCCGATACCTTCTTGGAGTTTCTACAG	518
QY	134	GlnLysGlyGlnLysLeuSer-GlnAspGlnThrValLeuLeuGly-ValAspPheProAs	153
DB	519	AGAAAGGGAAGAGTGTGAGCCAGAGTATAGAGGCTCTCCGCGAGCATGGATTTCCCTGA	578
QY	153	PserSerSerProSerCysThrLeuGlyLeuValLeuProLeuPserAspThrGlnVa	173
DB	579	CAGCAGCTCCCCAGCTCACCCTGGGCGCTGTCTTGGCCCTTGGAGTGCACACCCAGGT	638
QY	173	IlyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlnSerArgIlePheLy	193
DB	639	GTAATTATATGAGAGCGGGGCTTCAGGTGACGTCGTGGTGGCAAGCCGAGATCTTCAA	698
QY	193	SProIleSerIleGlnThrMetThrPalaThrLeuGlnValLeuHisGlnAlaCysGluAl	213
DB	699	GCCCATTCCTCATCCAGACCAATGTGGGCACACCTCAGTATTGCACACAGATCTGAGGC	758
QY	213	AlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrThrPalaSerHisTyrGl	233
DB	759	AGCTTAGGACAGCGGCTTGTACCGGGGAGAGTGCCTTCACTCGGGCCAGCCACTACCA	818
QY	233	ngluArgLeuAsnSerGlnGlnSerCysLeuAsnGlnThrPheAlaMetAlaAspLeuGl	253
DB	819	GGAGAGACTGAACCTCCGAACAGAGAGTGCCTCAATGAGTGAACGGCTATGGCCAGCTGGA	878
QY	253	userLeuArgProProSerAlaGluProGlyGlySerSerGlnGlnGlnMetGluGl	273
DB	879	GTCCTGGGCGCTCCACAGCGCGGAGCTGGCGGGTCTCTCAAAACAGAGAGATGAGGCG	938
QY	273	nalIleArgAlaGlnLeuThrPryValLeuAspValSerAspLeuIserValThrse	293
DB	939	GCGCATTCGCTGTGAGCTGTGAAAGTGTTCAGTGTGCTGATCTGGAGAGATGTGCGCTTC	998
QY	293	IlyGlnIleArgGlnAlaLeuGlnLeuArgLeuGlnValLeuProLeuGlnGlnIlyrArgAs	313

[illegible]

Pred. No. :	5, 51e-96	Length:	2061
Score:	1470.50	Matches:	290
Percent Similarity:	94.87%	Conservative:	6
Best Local Similarity:	92.95%	Mismatches:	15
Query Match:	60.37%	Indels:	1
DB:	6	Gaps:	1
US-09-761-640-4 (1-471) x AX405828 (1-2061)			
QY	1	MetAlaLeuValThrValSerArgSerProProGlySerClyAlaSerThrProValGly	20
Db	147	ATGGCCCTGGGACAGATGACCCGTTCCGCCCCCGGACGGGGGCGCTCCACGCCCTGGGG	207
QY	21	ProThrAspGlnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal	40
Db	207	CCCTGGGACCGGGGCTCCAGCAGAGAGAGTCCAGCTCCACGAGAGCAGAGCTTTGGGTG	267
QY	41	LeuArgGlyAlaValLeuGlyLeuGlnInAspGlyGlyAspAsnAspAlaAlaGlyAla	60
Db	267	CTCCGGTGGGGGTGCTCCGGGACTCCAGAGATGGAGGGGAGCAATGATATGACAGAGAGCC	327
QY	61	SerSerGluProThrGlnGlyAspAlaProSerGlyGlnGlnGlnGlyAspGlnThrAsp	80
Db	327	AGTTCTGAGCCACACAGAGAGAGCCCGCATGAGAGAGAGACTCCACGGGAGCCAGACAGC	386
QY	81	PheAlaGlnLeuSerGlnSerProGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	100
Db	387	TTGGGGCAAGGATCCAGAGATCCCGCCAGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG	446
QY	101	MetValGlnLeuLeuArgProGlnInAspAlaIleArgLeuAlaAlaGlnLeuGlnAlaPro	120
Db	447	ATGGTACAGCTGCTGAGGGCGCAGAGATGATCATCCCGCTGGACGCCACACTGGAGGACCC	506
QY	121	ArgProProArgLeuArgGlyLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSer	140
Db	507	CGGCTCTCCCGGGCTCCGACTCCCTGCTGGTAGTTTCTACAGAGAAAGAGAGGTCTGAGC	566
QY	141	GlnAspGlnThrValLeuLeuGlnGlyAlaAspPheProAspSerSerProSerCysThr	160
Db	567	CAGAGTAGAGCGGTCTCTGGGGCTGGATTTCCCTGACAGCAGACTCCCGACCTCGACC	626
QY	161	LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly	180
Db	627	CTGGGCTGTGTTCCCTCCCTCGAGGTGACACCCAGGATGACTAGATGAGAGCGGGGG	686
QY	181	PheSerValThrSerGlyGlySerArgGlyLeuPheLeuProIleSerIleGlnThrMet	200
Db	687	TTTCAGGTAAGCTCTGGTGGGCAAGGCGGATCTTTCAGGCCATCTCCATCCAGACATAG	746
QY	201	ThrAlaThrLeuGlnValLeuHisGlnAlaLeuGlyGlnAlaIleLeuGlySerGlyLeuVal	220
Db	747	TGGGGCACACTCCAGATGTTTCCACCAAGCATGTAGGACACTCTAGGCAAGCGGCTTGA	806
QY	221	ProGlyGlySerAlaLeuThrThrAlaSerHisIleGlnGlnArgLeuAsnSerGlnGln	240
Db	807	CCGGGTGGCAGTGGCTTACCTGGGGCCAGCCACACTACAGAGAGACTGAACTCCGACAG	866
QY	241	SerCysLeuAsnGlnTyrThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla	260
Db	867	AGCTCCCTCAATGATGATGACGGGTATGGCCAGCTGGAGTCTGGGGCTCCGACGCC	926
QY	261	GluProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	280
Db	927	GAGCTTGGCGGCTCTCAGACACAGAGACAGATGAGACAGCGCATCTCGTCTAGCTGTGG	986
QY	281	LysValLeuAspVal---SerAspLeuGlnSerValThrSerGlyGlnIleArgGlnAla	299
Db	987	AAAGTGTGGAGCTGAGAGACACTCAGAGACAGACAGATGACATGCGAGAGGTCTTCTCTCC	1046
QY	300	LeuGlnLeuArgLeuGlnGlyLeuProLeuGlnGlnGln 311	
Db	1047	CACGAGCTTTCACATGAGAGAGCGCTTGGACGCCCTTC 1082	

RESULT 10
AK000522 2064 bp mRNA linear PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20515 fis, clone KAT09889.
DEFINITION AK000522
ACCESSION AK000522.1 GI:7020673
VERSION oligo capping: fis (full insert sequence).
KEYWORDS Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone_lib:KAT clone:KAT09889.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Matanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
Unpublished
2 (bases 1 to 2064)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
TITLE Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source location/Qualifiers
1..2064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT09889"
/cell_line="KATO III"
/cell_type="signet-ring cell carcinoma"
/clone_id="KAT"
/note="cloning vector pME18SFL3"
147..1331
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91228.1"
/db_xref="GI:7020674"
translation="MALVTYSRSPSGASTPVGPMDOAVORSLQROSFVLRGA VLIOGCGDMDAAEASSEPEKAPSEELHGDTFGGOSPOKEORHLILMY OLIRPDIDILALQLEAPRRLRYLLIVSTRREGESLODEYILGVDPSSSPCT LGLVPLMSDTQYVLDGGDFSVYSGGOSRIFRPSIGTMMATLOYLHOCFAISG LVPFGSALTMASHYORLNSGCLNMTMADLESPRPSAPRGSSRDEKQEAR AELMKVLELSETSSETSDMPVESHSHESHEPPLQFPLARTGGQVDVGPALKS ROSVVTLOGSAVVANRTOAFQEOEGOGGOGEPCLISSTPRFRKVVKQASVHDSGEG EA"
BASE COUNT 421 a 671 c 605 g 367 t
ORIGIN
Alignment Scores:
Pred. No.: 5.52e-96 Length: 2064
Score: 1470.50 Matches: 290
Conservative: 6
Best Local Similarity: 94.87% Mismatches: 15
Query Match: 92.95% Indels: 1
DB: 60.37% Gaps: 1
US-09-761-640-4 (1-471) x AK000522 (1-2064)
QY 1 MetalLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
Db 147 ATGGCCCTGGTCAAGTGAAGCCGTTCCGCCCGGACGCGGCTCCACGCCCGTGGGG 206
QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40

|||||
Db 207 CCTGGGACGAGCGGCTCCACGAGAGTGCATCCAGCGAAGCGAGCTTGGCGGTG 265
QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGln 60
Db 267 CTCCTGGGCGGCTGCTGGGACTGCAGATGAGAGGACATGATGATGACGACGAGGCC 326
QY 61 SerSerGlnProTrpGlnValAspAlaProSerGlnGlnGlnGlnGlnGlnGlnGln 80
Db 327 AGTTCTGAGCCACAGAGAGAGGCCCCGAGTGAAGGAGAGCTCCACGGGAGCCAGAC 386
QY 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
Db 387 TTGGGCAAGAGATCCACAGATGCCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 446
QY 101 MetValGlnLeuLeuLeuArgProGlnAspAspAlaArgLeuAlaAlaGlnLeuGln 120
Db 447 ATGCTACAGCTGTGAGAGCCCGCAGATGACATCCGCTGGCAGCCCGCTGGAGCACC 506
QY 121 ArgProProArgLeuArgGlyTrpLeuValValSerThrArgGlnGlnGlnGlnGln 140
Db 507 CGGCTCCCGGCTCCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
QY 141 GlnAspGlnThrValLeuLeuGlnGlyValAspPheProAspSerSerProSerCysThr 160
Db 567 CAGATGAGAGAGCTCTCTGCGGCTGGATTTCCCTGACACACACTCCCGCAGCTGCACC 626
QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTrpLeuAspGlyAspGly 180
Db 627 CTGGGCTGTGCTTCCCTCTGTGAGTACACCAGAGGTACTTAAATGAGAGGGGGG 686
QY 181 PheSerValThrSerGlnGlnGlnSerArgGlnPheLeuProIleSerIleGlnThrMet 200
Db 687 TTCAGCGAGAGCTGTGTGGCAAGCGGATCTTCAAGCCATCTCATCAGACCATG 746
QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
Db 747 TGGCCACACTCCAGATATTGTCACCAAGCATGAGGAGAGCTTGAAGGAGGCGCTGTA 806
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisIleGlnGlnGlnGlnGlnGlnGln 240
Db 807 CCGGCTGGACAGTCCCTCCTGAGTACACCAGAGGTACTTAAATGAGAGGGGGG 866
QY 241 SerCysLeuAsnGlnTrpThrAlaMetAlaAspLeuGlnSerLeuArgProSerAla 260
Db 867 AGCTGCTCATAGTGTGAGCGGCTATGCGCCACTGAGTCTTGGCGCTCCACGCGCC 926
QY 261 GlnProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 927 GAGCTGGCGGCTCTCAGAACAGAGAGAGATGAGAGCGGATCCGCTGAGCTGTGG 986
QY 281 LysValLeuAspVal---SerAspLeuGlnSerValThrSerGlnGlnGlnGlnGln 299
Db 987 AAATGTGTGAGCGTGAAGAGCACTTCAGACAGTGCATGCCAGAGGCTCTCTTCC 1046
QY 300 LeuGlnLeuArgLeuGlyLeuProLeuGlnGlnGlnTrp 311
Db 1047 CACGAGCTTTCATGAGAGAGCTTTCGAGCCCTTC 1082
RESULT 11
LOCUS AP002776 160903 bp DNA linear HTG 17-JAN-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-126P21 map 11q, WORKING DRAFT
SEQUENCE 36 unordered pieces.
ACCESSION AP002776
VERSION AP002776.2 GI:12246853
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-126P21.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 160903)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 160,903 genomic DNA of 11q
Published Only in Database (2000)
2 (bases 1 to 160903)

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (10-JUN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22, Sphiro-chou, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
On Jan 16, 2001 this sequence version replaced g1.9188614.

Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

Project Information

Center project name: Humdraft1

Center clone name: RP11-126P21

Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 149950 bases at least Q40

Consensus quality: 153939 bases at least Q30

Consensus quality: 156144 bases at least Q20

Insert size: 157403; sum-of-contigs

Quality coverage: 8.42x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
36 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 12626 contig of 12626 bp in length
12727 24539 contig of 11813 bp in length
24640 33605 contig of 8966 bp in length
33706 45529 contig of 11824 bp in length
45630 51226 contig of 5597 bp in length
51327 58567 contig of 7241 bp in length
58668 65488 contig of 6821 bp in length
65589 72315 contig of 6727 bp in length
72416 78297 contig of 5882 bp in length
78398 84113 contig of 5716 bp in length
84214 90536 contig of 6323 bp in length
90537 90551 contig of 4415 bp in length
90552 95151 contig of 4132 bp in length
95152 99283 contig of 4132 bp in length
99284 99383 contig of 100 bp
99384 103658 contig of 4275 bp in length
103659 103758 contig of 100 bp
103759 107823 contig of 4065 bp in length
107824 107923 contig of 100 bp
107924 113135 contig of 5212 bp in length
113136 113235 contig of 100 bp
113236 117634 contig of 4399 bp in length
117635 117734 contig of 100 bp
117735 121241 contig of 3507 bp in length
121242 121341 contig of 100 bp
121342 124894 contig of 3553 bp in length
124895 124994 contig of 100 bp
124995 128257 contig of 3263 bp in length
128258 128357 contig of 100 bp
128358 131408 contig of 3051 bp in length
131409 131508 contig of 100 bp
131509 134212 contig of 2704 bp in length
134213 134312 contig of 100 bp
134313 136588 contig of 2276 bp in length
136589 136688 contig of 100 bp
136689 139497 contig of 2809 bp in length
139498 139597 contig of 100 bp
139598 142074 contig of 2477 bp in length
142075 142174 contig of 100 bp
142175 143858 contig of 1684 bp in length
143859 143958 contig of 100 bp
143959 146242 contig of 2284 bp in length
146243 146342 contig of 100 bp
146343 148167 contig of 1825 bp in length
148168 148267 contig of 100 bp
148268 150246 contig of 1979 bp in length
150247 150346 contig of 100 bp
150347 151876 contig of 1530 bp in length
151877 151976 contig of 100 bp
151977 154322 contig of 2246 bp in length
154323 154422 contig of 100 bp
154423 155669 contig of 1247 bp in length
155670 155769 contig of 100 bp

159723 160903 contig of 1181 bp in length.
NOTE: This is a 'working draft' sequence. It currently
consists of 36 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 12626: contig of 12626 bp in length
12627 12726: gap of 100 bp
12727 24539: contig of 11813 bp in length
24540 24639: gap of 100 bp
24640 33605: contig of 8966 bp in length
33606 33705: gap of 100 bp
33706 45529: contig of 11824 bp in length
45530 45629: gap of 100 bp
45630 51226: contig of 5597 bp in length
51227 51326: gap of 100 bp
51327 58567: contig of 7241 bp in length
58568 58667: gap of 100 bp
58668 65488: contig of 6821 bp in length
65489 65588: gap of 100 bp
65589 72315: contig of 6727 bp in length
72316 72415: gap of 100 bp
72416 78297: contig of 5882 bp in length
78298 78397: gap of 100 bp
78398 84113: contig of 5716 bp in length
84114 84213: gap of 100 bp
84214 90536: contig of 6323 bp in length
90537 90636: gap of 100 bp
90637 95051: contig of 4415 bp in length
95052 95151: gap of 100 bp
95152 99283: contig of 4132 bp in length
99284 99383: gap of 100 bp
99384 103658: contig of 4275 bp in length
103659 103758: gap of 100 bp
103759 107823: contig of 4065 bp in length
107824 107923: gap of 100 bp
107924 113135: contig of 5212 bp in length
113136 113235: gap of 100 bp
113236 117634: contig of 4399 bp in length
117635 117734: gap of 100 bp
117735 121241: contig of 3507 bp in length
121242 121341: gap of 100 bp
121342 124894: contig of 3553 bp in length
124895 124994: gap of 100 bp
124995 128257: contig of 3263 bp in length
128258 128357: gap of 100 bp
128358 131408: contig of 3051 bp in length
131409 131508: gap of 100 bp
131509 134212: contig of 2704 bp in length
134213 134312: gap of 100 bp
134313 136588: contig of 2276 bp in length
136589 136688: gap of 100 bp
136689 139497: contig of 2809 bp in length
139498 139597: gap of 100 bp
139598 142074: contig of 2477 bp in length
142075 142174: gap of 100 bp
142175 143858: contig of 1684 bp in length
143859 143958: gap of 100 bp
143959 146242: contig of 2284 bp in length
146243 146342: gap of 100 bp
146343 148167: contig of 1825 bp in length
148168 148267: gap of 100 bp
148268 150246: contig of 1979 bp in length
150247 150346: gap of 100 bp
150347 151876: contig of 1530 bp in length
151877 151976: gap of 100 bp
151977 154322: contig of 2246 bp in length
154323 154422: gap of 100 bp
154423 155669: contig of 1247 bp in length
155670 155769: gap of 100 bp

Db 86107 AAGTGACAGGTTAAAAAACAACAAAAAAACTGGCGGGGCGGTTGGCTCATGGC 86166
Oy 113 ----- 113
Db 86167 TGTATCCAGCATTCTGGAGGCCGAGCGGAGATCAGAGTCAAGAGATCGAGAC 86226
Oy 113 ----- 113
Db 86227 CATCGGCTAACAGCGTGAACCCCGTGTCTACTAAAAACAAAAAATTAGCCAGCGG 86286
Oy 113 ----- 113
Db 86287 TGTGGCGGGTCTGTAGTCCAGCTACTCGGAGGCTGAGGACAGAGAATGCGTGA 86346
Oy 113 ----- 113
Db 86347 CCGGGAGATGAGATTGTAGTAGCTGAGATTGCCCACTCCAGCTCCAGCCGGGTGAC 86406
Oy 113 ----- 113
Db 86407 AGAGCAAGATCCATTGCAAAAAAAGAAAAAACCTCCTTCTTGGGGGCACGT 86466
Oy 113 ----- 113
Db 86467 GGTTCACGCTGTAAATCCAGCACTTTGGAGGCTGAGAGGAGGATTGCCAGTCCA 86526
Oy 113 ----- 113
Db 86527 GGAGTTCAAGACCCAGCCCAAGAACATAGGAGAGATCTCATCTCAACAAAAAATTAAAA 86586
Oy 113 ----- 113
Db 86587 ATTATCTGGGATGTGTGCATGGCGCTGTATGCCAGCACTTGGGAGGCTAAGTGGA 86646
Oy 113 ----- 113
Db 86647 GGATCATGTAGCCTCGGAGATTGGGGCTACAGTAGCTGTGATGGCCACTGACTCC 86706
Oy 113 ----- 113
Db 86707 AGCTGGGCGACAGTAGAGACCTATCTCAAAACAAAAAACAACCTCTTCTC 86766
Oy 113 ----- 113
Db 86767 AGCAGAAACAAAAATGTGCAACTATGTCACTCCATCCCGCTTTTTCACACAGT 86826
Oy 113 ----- 113
Db 86827 CCCTCTTACCCACTGATTTTACATCTGGGAAGAGGGGGATGGAGAGAGAGACAG 86886
Oy 113 ----- 113
Db 86887 GTCAAGACTGTTCTGGGCTGGGGTGTAGGTGGGAGCAGGGTCTTGGGCCACTCT 86946
Oy 114 ----- 124
Db 86947 GACCTGGGTTCTCATCTCCCGCCAGCGCCAGCTGGAGGACCCCGGCTCCCGG 87006
Oy 125 ----- 144
Db 87007 CTCCGCTACCTGCTGTGTAGTTTCTACAGAGAGAGAGAGAGGCTGAGCCAGATGAGACG 87066
Oy 145 ----- 154
Db 87067 GTCCCTCGGGCGTGGATTCTCCCTGACAGCATTCGACAGAGGAGAGAAAGAGGGGCG 87126
Oy 155 ----- 158
Db 87127 AAAGAGTGAGAGAGAGGGGAGACAATGCTTCCCTCTGCCCCCTGCAGGTCCCCAG 87186
Oy 158 ----- 178
rcysThrLeuGlyLeuValLeuProLeuTrpSerAspHrGlnValTrpLeuAspGlyAs 178
|||||

Db 87187 CTGCACCTGGGGCTGTGTCTTGCCCTTGAGTGCACCCAGGTACTTATGATGAGA 87246
Oy 178 ----- 179
Db 87247 CGG-CTAACCATATGCCACTGAGAGCTGGGGGCTGAGAGGTGGAATACTAGTGTACGG 87305
Oy 179 ----- 179
Db 87306 ATGTGACGGGGCTGGGACACAGGTCAGTAGAACTTAGGCTGGAGCTGGGGCTTG 87365
Oy 179 ----- 179
Db 87366 CTCCAGCTCCACGTGAGATTACCGTATCCGAGCAAGGACCTGACTCCTGGGTCTC 87425
Oy 179 ----- 179
Db 87426 GGCTTCCGTATTCAGCAAAAGATGGTTCTGTGCGAGAGCCTGAGTCTGGCTCA 87485
Oy 180 ----- 192
Db 87486 CCTGTCCCTGTCTCCTGCGAGGGGCTTCAGCGTGAAGTGTGGTGGCAAGCCGATCTT 87545
Oy 192 ----- 201
Db 87546 CAAGCCATCTCCATTCAGACCATGTGTAAGACAGACAGACTGCTGACATCAGGCCAGC 87605
Oy 201 ----- 201
Db 87606 CTCTCCTTCGAAGAGGCTGCAAGGCTGGGAGGCGACGAAGGAAACAGTGAACCA 87665
Oy 202 ----- 215
Db 87666 GCCTCTCCTCTCTCTCTCAGAGGCGACACTCCAGGTTGTGACCACAGATGTGAGCACTC 87725
Oy 215 ----- 235
Db 87726 TAGCGACGGCCTGTACCGGCTGGCAGGCTCCATCCCTGAGCCAGCCATCCAGGAGA 87785
Oy 235 ----- 255
Db 87786 GACTGAACTCCGACACAGAGCTCTCAATGAGTGAAGGAGGAGGAGTGGAGTCTC 87845
Oy 255 ----- 264
Db 87846 TGGCGCTGCCAGGCGCCAGCCTGGCGGGTCAGTGTGTGGAGGGAGGAGTGGGTGAG 87905
Oy 264 ----- 264
Db 87906 GCGAAGCAGATATATGACAGTGGGGGCGCATGTGGAGAGAGGAGCCCTTGACTGA 87965
Oy 265 ----- 279
Db 87966 GGGGCTGTGCTCCAGGTCTCTCAGAACAGAGCAATGAGAGGCGATTCGCTGTAGC 88025
Oy 279 ----- 295
Db 88026 TGTGAAAGTGTGTGATGACGTGAGAGTGTGACATTCCTCCAAAGAGGTGGGCGAGG 88085
Oy 295 ----- 295
Db 88086 GGCCGGGAGTGAAGTCCCTTACAGAGGGCTGCACAGCTTGCCTTCTGGAGCCCTC 88145
Oy 295 ----- 295
Db 88146 CCCACTTGTCTGTCTCCAGTCTGAGCCATTCTGATGCTTGGCTTGGCCCTAATC 88205
Oy 295 ----- 295
Db 88206 CATCTAATGGGATAGGGTGTGTCTCAGAGGTCCAGAGGAAGTGGCAGGTGGCG 88265
Oy 296 ----- 301
Db 88266 ACTAGAGAGCCCGAGCCTCTTCTGCTTGGGCTGTGCTGCTGCTGAGATCGCGCAGGCTCTGG 88325
|||||

QY	301	luleuArlgLeuGllyleuProLeuGInGIntYrArAspheIlleAspaSngInMeLleuL	321
Db	88326	AGCTGGCCCTGGGGCTCCCTCCACGACGTACCGTGAATTATCCACAACCATGATGTGC	88395
QY	321	euleuValAlaGlnrArqAsparAlaSerArgIlePheProHisLeuYrIleu-----	338
Db	88386	TGCTGGTGGCACAGGGGGACCGAGCGCTCCCGATCTTCCCCACCTTAACCTGTGAGCT	88445
QY	338	-----	338
Db	88446	TCAGCCAGGCGCTGGGCCATGGGGACCGCTGGCAGCTGTAGAGGGCGGACGGCCAGAGCAG	88505
QY	339	-----GlySerGlnTrpAsnIle	344
Db	88506	AGCGTGAAGAGAGAGAGAAAATGTGCTGTTCCCTCTCCACAGGGCTCAAGTGAACGCA	88565
QY	345	AlaAsnLeuGlnGluLeuGlnArqAsnArqVal-----	355
Db	88566	GCAAACCTGGAGAGAGCTGCAGAGAAACAGTA - GGGCATATGAGCCCTCGGGCCACCCAC	88624
QY	355	-----	355
Db	88625	CCCATCTTCCCTTCCTGCTGCGCTCCCGCATTTGGTGTAGCCAGCTTCAAAACCCCTG	88684
QY	355	-----	355
Db	88685	GACCACTCAGACAGCTGCTAGCTGCTTCTTAACCTGTCTCTGGGGCTGTGGCCCTGGT	88744
QY	355	-----	355
Db	88745	GTGGGCTCCAGAGTGGGACAGAGACACTGTGGCCAGCCCCCGGCCACTCTCCCTCCC	88804
QY	355	-----	355
Db	88805	ATCCACACTGTGAACAAGACAGACAAACAAGGGCCTAGCCACGCCAAGACGAGAACA	88864
QY	355	-----	355
Db	88865	GCAGGCATACTGCTGTAACTGCTTTGGACAAGAGAAAAGGCTCTCTGTGAATGCGCC	88924
QY	355	-----	355
Db	88925	TGTGGGCCAGCTACTTGGGAGGTGAGGCGACGAGAGATGCTTGAGCCCTGGAGATTGAG	88984
QY	355	-----	355
Db	88985	GCCGAGTGAGCCGTGATCAGCGCCACTGCACCTCGAGCCTGGGCAACAGAGAGACCTG	89044
QY	355	-----	355
Db	89045	TCTTAAAAAATTAAGAAAAAAGAGAGAGAAAAAGCCTTTCTCTCACCTTCCCTGTCT	89104
QY	355	-----	355
Db	89105	CAGGAGAGAGAACTGCCCTTCTCCCGTGGGACCTGGCTGCTGTGACAGGTAC	89164
QY	355	-----	355
Db	89165	CTGTATCTGCCACCATGGGGTTCTGGGACCTGTGTAGCCCTGCAACCCACCTGCTGC	89224
QY	355	-----	355
Db	89225	AGACCCACCCACTTCAGCTTAGCTCAAAAGCTGTCTTAACATCTTGTGAGAAATAAC	89284
QY	355	-----	355
Db	89285	TGAAGGCTGGAGTTGCAGTTGGCCACAGTGTCTGGACCAGATGGGGAAACAAGCCAGC	89344
QY	355	-----	355
Db	89345	AGGGCAAGTGAATTGTCTAAAGTGCAGCCAGCGTGCAGCTGGGTCTACTTCTCTCCA	89404

QY	355	-----	Thrs11leuansmeta1Ar9glu1leasPasnPhetYrProglu	371	355
Db	89405	CTGTCACTGCTGCGCCATCTGACTTGAACATCAGAGCTAAAGCGAGGCCAGGGCTGC			89464
QY	355	-----			355
Db	89465	CAGGTCCGATTGCCAGTGGGCTGGTACCTTGCACCATACAGCTGGCCTTGGGCGCTC			89524
QY	356	-----	Thrs11leuansmeta1Ar9glu1leasPasnPhetYrProglu	371	
Db	89525	TGGGACAGGGGTACCCCATCTTTGAACATTGGCGCGGAGATTGACAACTTCTACCCCTAG			89584
QY	372	ArgPhetYrThYrHisAsnValArgLeuTrpAspGluJuseRalaGlnLeuLeuProHis			391
Db	89585	CGCTTCACCTACCAACAAATGTGGCCTCTGGGATGAGAGTGGGCCCACTGCTGCCGAC			89644
QY	392	TrpYrSglUthrHisArgPheIleGluIalaIa			402
Db	89645	TGGAAGGAGACGACCGCTTCATGTAGGCTGC -AAGTCGGCTTCCTGCTTGGCTGCT			89703
QY	402	-----			402
Db	89704	GCCCCACGTGGCTTCAGAGCGGCGCTGCCCTCCCTCCCTGCCCCCATCTGCACACT			89763
QY	402	-----			402
Db	89764	GTGCATTCCTTCCAGCCCTCAGTGTCTCCCTCTCTCTACTCTCAGTTGTTGTA			89823
QY	402	-----			402
Db	89824	GCCCTTCACCTGCCCTCCCTGACAGAGCGGGATCCTTCATGTGAGGGAGGCTGCTG			89883
QY	403	-----	ArgIaGlnGlyThrHisValLeuValHisCysIysMetGly		416
Db	89884	AACCTGGCGCTGTTCACACAGACACAGGGCACCCACGCTGTGTCTCACTGCAGAATGGGC			89943
QY	417	ValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGluCysSerLeu			436
Db	89944	GTTCACCGCTCAGCGGCACACAGTGTGGCTATGCGATGAAGCAGTACGAATGCAGCTTG			90003
QY	437	GluGlnIaLeuArgHisValGlnIuLeuArgProIleAlaArgProAsnProGlyIhe			456
Db	90004	GAGCAGGCCCTGCGCAGCTGCAGGAGGCTCGCGGCCCATCGCCGCCCAACCTGGCTTC			90063
QY	457	LeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAlaArg			470
Db	90064	CTGCGGCAGCTGACATCTACAGGGGCATCTCGACGGCCAGG			90105
RESULT 12					
AP001885		171566 bp	DNA	Linear	HTG 13-JAN-2001
LOCUS					
DEFINITION		Homo sapiens chromosome 11 clone Rpl1-157K17 map 11q, WORKING DRAFT			
SEQUENCE		24 unordered pieces.			
ACCESSION		AP001885			
VERSION		AP001885.3			
KEYWORDS		HTG; HTGS; PHASE1; HTGS; DRAFT.			
SOURCE		Homo sapiens DNA, clone:Rpl1-157K17.			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE		1 (bases 1 to 171566)			
JOURNAL		Hatori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,			
REFERENCE		Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.			
AUTHORS		2 (bases 1 to 171566)			
TITLE		Published Only in Database (2000)			
JOURNAL		Homo sapiens 171,566 genomic DNA of 11q			
REFERENCE		1-7-22 Suehito-chou, Tsuromi-ku, Yokohama, Kanagawa 230-0045, Japan			
AUTHORS		(E-mail: hatori@rics.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/			
JOURNAL		Submitted (24-APR-2000) Mesahira Hatori, The Institute of Physical			
		and Chemical Research (RIKEN), Genomic Sciences Center (GSC);			
		1-7-22 Suehito-chou, Tsuromi-ku, Yokohama, Kanagawa 230-0045, Japan			
		(E-mail: hatori@rics.riken.go.jp, URL: http://hnp.gsc.riken.go.jp/			

COMMENT

Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 13, 2001 this sequence version replaced gi:8117536.

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: <http://hgp.gsc.riken.go.jp/>

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: HumDrafit1

Center clone name: RP11-157K17

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 164093 bases at least Q40

Consensus quality: 166550 bases at least Q30

Consensus quality: 167995 bases at least Q20

Insert size: 169266; sum-of-contigs

Quality coverage: 8.92x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1      15114 contig of 15114 bp in length
15215 27885 contig of 12671 bp in length
27986 42419 contig of 14434 bp in length
42520 55000 contig of 12481 bp in length
55101 68523 contig of 13423 bp in length
68624 79107 contig of 10484 bp in length
79208 89207 contig of 10000 bp in length
89308 95794 contig of 6487 bp in length
95895 103725 contig of 7831 bp in length
103826 119213 contig of 8280 bp in length
119314 124281 contig of 7008 bp in length
124382 130027 contig of 5646 bp in length
130128 135012 contig of 4885 bp in length
135113 138400 contig of 3288 bp in length
138501 144921 contig of 6421 bp in length
145022 150481 contig of 5459 bp in length
150581 153837 contig of 3257 bp in length
153938 156948 contig of 3011 bp in length
156949 157048 contig of 3011 bp in length
157049 161384 contig of 4336 bp in length
161385 161484 contig of 100 bp in length
161485 164413 contig of 2929 bp in length
164514 167396 contig of 2883 bp in length
167497 169991 contig of 2495 bp in length
170092 171566 contig of 1475 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1      15114 contig of 15114 bp in length
15115 15214 gap of 100 bp
15215 27885 contig of 12671 bp in length
27886 27985 gap of 100 bp
27986 42419 contig of 14434 bp in length
42420 42519 gap of 100 bp
42520 55000 contig of 12481 bp in length
55001 55100 gap of 100 bp
55101 68523 contig of 13423 bp in length
68524 68623 gap of 100 bp
68624 79107 contig of 10484 bp in length
79108 79207 gap of 100 bp
79208 89207 contig of 10000 bp in length
89208 89307 gap of 100 bp

```

FEATURES

source

```

*      89308      95794: contig of 6487 bp in length
*      95795 95894: gap of 100 bp
*      95895 103725: contig of 7831 bp in length
*      103726 103825: gap of 100 bp
*      103826 112105: contig of 8280 bp in length
*      112106 112205: gap of 100 bp
*      112206 119213: contig of 7008 bp in length
*      119214 119313: gap of 100 bp
*      119314 124281: contig of 4968 bp in length
*      124282 124381: gap of 100 bp
*      124382 130027: contig of 5646 bp in length
*      130028 130127: gap of 100 bp
*      130128 135012: contig of 4885 bp in length
*      135013 135112: gap of 100 bp
*      135113 138400: contig of 3288 bp in length
*      138401 138500: gap of 100 bp
*      138501 144921: contig of 6421 bp in length
*      144922 145021: gap of 100 bp
*      145022 150480: contig of 5459 bp in length
*      150481 150580: gap of 100 bp
*      150581 153837: contig of 3257 bp in length
*      153838 153937: gap of 100 bp
*      153938 156948: contig of 3011 bp in length
*      156949 157048: gap of 100 bp
*      157049 161384: contig of 4336 bp in length
*      161385 161484: gap of 100 bp
*      161485 164413: contig of 2929 bp in length
*      164414 164513: gap of 100 bp
*      164514 167396: contig of 2883 bp in length
*      167397 167496: gap of 100 bp
*      167497 169991: contig of 2495 bp in length
*      169992 170091: gap of 100 bp
*      170092 171566: contig of 1475 bp in length.

FEATURES
source
1. 171566
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q"
   /clone="RP11-157K17"
1. 15114
   /note="assembly-fragment clone_end:T7 vector_side:right"
15215..27885
   /note="assembly-fragment"
27986..42419
   /note="assembly-fragment"
42520..55000
   /note="assembly-fragment"
55101..68523
   /note="assembly-fragment"
68624..79107
   /note="assembly-fragment"
79208..89207
   /note="assembly-fragment"
89308..95794
   /note="assembly-fragment"
95895..103725
   /note="assembly-fragment"
103826..112105
   /note="assembly-fragment"
112206..119213
   /note="assembly-fragment"
119314..124281
   /note="assembly-fragment"
124382..130027
   /note="assembly-fragment"
130128..135012
   /note="assembly-fragment"
135113..138400
   /note="assembly-fragment clone_end:SP6 vector_side:left"
138501..144921
   /note="assembly-fragment"
145022..150480
   /note="assembly-fragment"

```



```
/note="assembly_fragment"
misc_feature 150581..153837
/note="assembly_fragment"
misc_feature 153938..156948
/note="assembly_fragment"
misc_feature 157049..161384
/note="assembly_fragment"
misc_feature 161485..164413
/note="assembly_fragment"
misc_feature 164514..167396
/note="assembly_fragment"
misc_feature 167497..169991
/note="assembly_fragment"
misc_feature 170092..171566
/note="assembly_fragment"
BASE COUNT 45656 a 39885 c 40004 g 43721 t 2300 others
ORIGIN
Alignment Scores:
Pred. No.: 5,54e-89 Length: 171566
Score: 1405.50 Matches: 445
Percent Similarity: 24.94% Conservative: 3
Best Local Similarity: 24.78% Mismatches: 13
Query Match: 57.70% Indels: 1338
Gaps: 13
US-09-761-640-4 (1-471) x AP001885 (1-171566)
OY 10 ProProGlySerGly-----AlaSerThr----- 17
Db 201 CCAAGGGGCCCGGTTAGAAATGACAGATGCTCAGGCTCCACTAGACCTATCGACC 260
OY 18 -----ProValGlyProThrPaspGlnAlaValGln-----ArgArgSerArg 31
Db 261 GCGTGGGGCCCATGAATCCGTGTAGACAAAGGCTGCAGGTGAATGTGATGCATGAGA 320
OY 32 Leu-----GlnArg 34
Db 321 CTTTGAGAGCCCTACCCCTAAGGGGAATGAGAGCCATGACAGTTGCCCATCTGCCCTGCC 380
OY 35 ArgGlnSerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsn 54
Db 381 AGGAGAGACCTTGGCGGTCTCCGTGGGGCTGTCTCGGAGTCAGAGATGAGAGGGACAAAT 440
OY 55 AspaAlaAlaGlnAlaSerSerGlnProThrGlnLysAlaProSerGlnGlnGlnLeu 74
Db 441 GATGATGACAGCAGAGGCGAGTTCTGAGCCACAGAGAGGCCCGCAGTGAAGAGAGCTC 500
OY 75 HisGlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnIlyGlnGlnGln 94
Db 501 CACGGGGACACAGACAGACTTCGGCCAAAGATCCCGACAGTCCCGAAGACAGAGAGAGCAG 560
OY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeu--- 113
Db 561 AGGCAGACCTGCACCTCATGTGTACAGCTGCTGAGGCGCCAGAGATGACATCCGCTGGTG 620
OY 113 ----- 113
Db 621 AGGGCCCATGTGAGCTCCGGGGGTGGGGGAAGACACACCTGAGGCGAGAAATGAG 680
OY 113 ----- 113
Db 681 CCCTGTGATGTGTCTGCTGGGAGCAATGTTGAGAGGCGAGGGGCGAGCTGGGAGT 740
OY 113 ----- 113
Db 741 GTTGGGTTACCATATCCAGTGCCTGCTGTTTCAGGACAGAGATGTTTGTGTTTGG 800
OY 113 ----- 113
Db 801 TTTGTTTTTTTTTTTGTGAGATGAGTCTGCTGTGTTGCCAGGCTGGAGTGGCGCTG 860
OY 113 ----- 113
```

```
Db 861 ATCTGGCTCACTGCAGAGTTCACACTCCCGGGTTACACCAATTCCTGCTCAGACCTCC 920
OY 113 ----- 113
Db 921 CGAGAGCTGGAGCTACAGGACCCGCCACTATGCTGGCTAATTTTTTTTGTATTTTGA 980
OY 113 ----- 113
Db 981 GTAGAGATGCGGTTTCACCCGTGTTAGCCAGGGTGTCTCAATCTCTGACCTCGTATCC 1040
OY 113 ----- 113
Db 1041 GCTTGCCCTCGGCTCCCTAAGTGTGGGATTACAGGCGTGAGCCACGCGCCGCGCTGG 1100
OY 113 ----- 113
Db 1101 GCAGAGATATTTTGCACAGGCGGAGGTGATGCTGACAGTCTCTGGAAAGAAAGACAG 1160
OY 113 ----- 113
Db 1161 GGAGGTTTGTCTCTGTGTGAGCCAAATAGGGCCAAATTAGTCCAAAGGGGTGGAAATTCAG 1220
OY 113 ----- 113
Db 1221 AGGCATGTTCTCCTCAACAGTAGAAGACTTTCCTGATGATGTAGCTGTCCCTAATG 1280
OY 113 ----- 113
Db 1281 GACCATGTCACAGGACAGGCTTGTGTGACACACACTGGGGGAGGGTGTGTAGAGGACC 1340
OY 113 ----- 113
Db 1341 CAGGACAGGGGAGAGGAGACTCCAGCCCAAGAGAAAGGATGCCAGTGGAGCTGAG 1400
OY 113 ----- 113
Db 1401 CGCTCATCTCAATATCTTGTGTGAGACAGCTGCGGCTCTCAGCAAGACGTGGCAACAA 1460
OY 113 ----- 113
Db 1461 GATCAGTGAATCTTGTGTAACACTGAATTTACAAAGAAAGTATCCTATCTGTGAG 1520
OY 113 ----- 113
Db 1521 ATTTGTCTTAATATTTCTTGTGTATAAACATTTCTCAATGAAATGATGATTTGCC 1580
OY 113 ----- 113
Db 1581 AAGTGACAGGCTAATAAAAAAAAAAACAACAAAAAAACTGGCGGGCGGGTGGCTCATGCC 1640
OY 113 ----- 113
Db 1641 TGTATCCCACTTGGGAGGCGGAGCGCGGAGATCAGAGAGTACAGAGATCGAGAC 1700
OY 113 ----- 113
Db 1701 CATCTGCTTAACACGGTGAACCCCGTGTCTACTAAAAACACAAAAAATTAGCCAGGCG 1760
OY 113 ----- 113
Db 1761 TGTGGCGGGTGCCTGTAGTCCAGACTACTCGGAGGCTGAGGACAGAGAAATGGGCTGA 1820
OY 113 ----- 113
Db 1821 CCGGAGATGAGAGTTGCACTGAGCTGAGATTTGGCCACTCCAGCCTGGGTGAC 1880
OY 113 ----- 113
Db 1881 AGAGCAAGCTCCATTTGCCAAAAAAGAAAAAAGAAAAAAGAACTCTTTCTTGTGGCACAGT 1940
OY 113 ----- 113
```

Db 1941 GGCACGCGCTTAATCCACGACTTTGGAGGCTGAGAAAGGAGATGCTGAGTCCA 2000
QY 113 ----- 113
Db 2001 GGACTTCAGACGACGCCAGAAACATAGGAAAGATCTCATCTCAACAAAACCTTAAA 2060
QY 113 ----- 113
Db 2061 AATTATCTGGGATGTTGGCATGGGCTGATGTCACGACAACTTGGAGGCTAAGTGGGA 2120
QY 113 ----- 113
Db 2121 GGATCATGTAGAGCTCGGAGTTGGGGCTACAGTGAAGCTGATGTGCCACTGTACTCC 2180
QY 113 ----- 113
Db 2181 AGCTGGGCGACAGATGAGAGCCCTATCTCAAAACAAAACAAACACCTCTTCTC 2240
QY 113 ----- 113
Db 2241 AGCAGAAACAAAATGTGCACCTATGTAGTCCATTCGCCGCTTTTTCACACAGT 2300
QY 113 ----- 113
Db 2301 CCTCTTTACCCACTGATTTTACATCTGGAGAGAGGGGATGGAGAGAGAGCAGG 2360
QY 113 ----- 113
Db 2361 GTACAGACTGTTCTGGGCTGGGGTGTAGTGGGAGCAGGCTCTTGGGCGCACTGT 2420
QY 114 ----- 114
Db 2421 GACCCTGGGTTCTCATCTCCGCCAGGACGACGCTGGAGGACCCGCGCTCCCGG 2480
QY 125 LeuArgTyrLeuLeuValSerThrArgGluGluGluSerGlnAspGluThr 144
Db 2481 CTCCGCTACCTGCTGTGTCTTACACAGAAAGAGAGGTGTGAGCCAGATGAGACG 2540
QY 145 ValLeuLeuGlyValAspPheProAspSer----- 154
Db 2541 GTCTCTGGGCGGTGATTTCCCTGACAGACGATTCGAGCAGGAGAGAAAGAGGGCG 2600
QY 155 ----- 155
Db 2601 AAAGAGGTGAGGAGACAGGGGGACAAATGCTTCCCTGCTCCCTCCAGCTCCAG 2660
QY 158 rCysThrLeuGlyLeuValLeuProLeuTyrPheSerAspThrGlnValTyrLeuAspGlyAs 178
Db 2661 CTGCACCTGGGCTGTGCTTGGCCCTGTGGAGTACACCCAGGTGTACTTATGAGAGA 2720
QY 178 pGly----- 179
Db 2721 CGG-GTAAGCAATGGAACACTGAGGTGGGGCTGGAGGTGGAATTACTGATGTGACGG 2779
QY 179 ----- 179
Db 2780 ATGTGAGGGGCTGGGACACGAGGTACAGTGAACCTTATAGGCTGGGACTGTGGGCGCTG 2839
QY 179 ----- 179
Db 2840 CTCCACAGCTCCACGTCAGATTACCGTGAATCTGAGCAAGGACCTGATCTCGGGTCTC 2899
QY 179 ----- 179
Db 2900 GGCCTCCGATTCACAGCAAAAGATGGTCTCTGTGGCAGAGCCTGGAGTCTGGCCTCA 2959
QY 180 ----- 180
Db 2960 CCTGTGCTGTCTCTGCAAGGCTTCACAGCTGACGTCTGTGGGAAAGCCGATCTT 3019
QY 192 eLysProIleSerIleGlnThrMetTrp----- 201
Db 3020 CAAGCCCATCTCATCCAGACCATGTGTAGAGACAGAGACTGCTGAGCTCAAGCCAGC 3079

QY 201 ----- 201
Db 3080 CTCTCCTTGAAAGAGAGCTGCAGGGCCTGGGGAAAGGACGAAGGGAACAGTGTGACCA 3139
QY 202 ----- 202
Db 3140 GCCTCTCTCTGCTCTCAGAGGCGCACCTCAGGTATGTGCCAACGATGTGAGGACGCTC 3199
QY 215 euGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluA 235
Db 3200 TAGGACAGGGCTTGTACCGGGTGGCAGTGCCCTCACTGGGCCACCCACTACAGAGAGA 3259
QY 235 rGluAsnSerGluGlnSerCysLeuAsnGluTPrThrAlaMetAlaAspLeuGluSerL 255
Db 3260 GACTGAACCTCCGAACAGAGCTGCTCAATGAGTGCAGGCGCTATGGCGCAGCTGAGTCTC 3319
QY 255 euArgProProSerAlaGluProGlyGly----- 264
Db 3320 TGGGCTCTCCACGCGGAGCCGTGGGCTCAGTGTGTGAGAGGAGGAGCTGGGTGAG 3379
QY 264 ----- 264
Db 3380 GGAAGGACAGATTAATGCAGTGGGGGCAATGGTGAAGAGGAAAGGCCCTTGACTGA 3439
QY 265 ----- 265
Db 3440 GGGGCTGTGCTCCAGGTCTCTCAGAACAGAGAGAGATGAGCAGAGGAGATCCGTGCTGAGC 3499
QY 279 euTrpLysValLeuAspValSerAspLeuGluSerValThrSerLysGlu----- 295
Db 3500 TGTGGAAGTGTGTGATGTACAGTACCTGAGAGTGTCACTTCCAAAGAGTGGGCGAGGG 3559
QY 295 ----- 295
Db 3560 GGGCGGGGAGTACGTCCTCTAGCAGGGGCTGCAGACTTGCTTCTGGAGCCCTC 3619
QY 295 ----- 295
Db 3620 CCCACTGTGCTGTCCAGTCTCTGAGCATTCCTGATGCTTGCGCTTGGCCCTAATC 3679
QY 295 ----- 295
Db 3680 CATCTAGATGGATAGGTCCTCTCTCAGAGTCCAGAGGAAAGTGGCAGATTGGGC 3739
QY 296 ----- 296
Db 3740 ACTAGAGAGACCCCAAGCTCTTGGCTTGGGCTGTGCTGCAGATCCGCGAGGCTTGG 3799
QY 301 LuLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuL 321
Db 3800 AGCTGGCGCTGGGGCTGCCCTCCACAGCATACCGTACATTCAACGACAGATGCTGC 3859
QY 321 euLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeu----- 338
Db 3860 TGCCTGTGGACACGCGGAGCCGAGCTCCCGCATCTTCCCCACACTTACCTGCTGAGCT 3919
QY 338 ----- 338
Db 3920 TCAGCCAGGCTGGGCGCATAGGGAGCCGTGGCAGCTGTAGAGGGCGGACGACGACGAG 3979
QY 339 ----- 339
Db 3980 AGGCTGGAGAGAGGAATGTGCTTCCCTCTCTCCAGAGGCTCAAGAGTGAAGACGA 4039
QY 345 AlaAsnLeuGluGluLeuGlnArgAsnArgVal----- 355
Db 4040 GCAAACTGTGAGAGCTGCAGAGAGAACAGTGA-GGGCTATGAGCCCTTGGGGCCACCCAGC 4098
QY 355 ----- 355
Db 4099 CCCATCTTCCCTTCTCTGCGCTCCCGCATTTGGGTGTAGCCAGCTTCAAAAACCCCTG 4158

QY 355 ----- 355
Db 4159 GACCACCTCAGCAGCTGCTAGCTCTGCTTCTACTCTGTCTGCGGGCTGTTGCCCTGCT 4218
QY 355 ----- 355.
Db 4219 GTGGGCTCCAGGTGGGGACAGAGACCTGTGCGCCGCCGCCACTCTCTCTCCCCC 4278
QY 355 ----- 355
Db 4279 ATCCACACTGTGAAACAGAGACAAGAGCCCTCAGCCAGCCCAAGACAGAGACA 4338
QY 355 ----- 355
Db 4339 GCAGGCATCTGCTGTACTGCTGTGACAGAGAAAAGGCTCTCTGTGATGCGCC 4398
QY 355 ----- 355
Db 4399 TGTGGGCCAGCTACTGTGGAGGCTGAGGACAGAGATGCTTGAGCCCTGGAGATTGAG 4458
QY 355 ----- 355
Db 4459 GCCGCACTGAGCCGTGATCAGCCACTGCATCTCAGCCTGGGCAACAGAGAGACCTG 4518
QY 355 ----- 355
Db 4519 TCTTAATAAATAAGAAAAAGAGAAAAAGCCCTTTCTCCACTTGCCTGTCT 4578
QY 355 ----- 355
Db 4579 CAGGGAAGAAGAACTGCCCTTCTCCCGTGGGACCTGGCTGCTGTGACAGATTAC 4638
QY 355 ----- 355
Db 4639 CTGTATCTGCCACCATGGGCTTGGAGCTGCTGTAGCCCTGCCACCCACTGCTGC 4698
QY 355 ----- 355
Db 4699 AGACCCACCTCTCAGCTTAGCTCAAAAAGCTTCTTAAGTCAATTTCTGAGATAAC 4758
QY 355 ----- 355
Db 4759 TGAAGGCTGAGTTGCAGTGGCCAGCTGTCTGACCATGAGAGAAACAAGCCACG 4818
QY 355 ----- 355
Db 4819 AGGCAAGATGATTGGTCTAAGGTGCGACGACGAGTGAAGTGGGTCACTTCTCTCCA 4878
QY 355 ----- 355
Db 4879 CTGTACTGTGCTCCATCTGACTTGAACATCAGAGCTAAAGCGAGGCCAGGGCTGC 4938
QY 355 ----- 355
Db 4939 CAGTTCGATTGCCAGTGGGCTGTACTCTGCCACATCAGCTGGGCTTGGGCGCTC 4998
QY 356 ----- 371
Db 4999 TGGGACAGGCTTACCCACATCTTGAACATGAGCCGGGAGATTGACAACCTTACCCCTGAG 5058
QY 372 ATGPhetThTyTHisasnValArgLeuTrpAspGluSerAlaGlnLeuLeuProHis 391
Db 5059 CGCTTACCTTACCAATGTGGCCCTCTGGATGAGAGACTCGGCCAGCTGCTGCCGCAC 5118
QY 392 TTPPLVSGlufThHisArgPheIleGluAlaAa----- 402
Db 5119 TGAAGAGACGACACCGCTTCAATGAGCTGC-AAAGTGGTCTCTCTGCTTGTGCTCT 5177
QY 402 ----- 402
Db 5178 GCCCCACTGTGCTTCAAGCGGCTCCGGTGCCTCTCTGCCACCTGTGCCATCAGCT 5237
QY 402 ----- 402

Db 5238 GTGCATTCTTCCAGCCCTCAGTGTCTTCCCTCTCTACTCTCAGTTGTTGA 5297
QY 402 ----- 402
Db 5298 GCCCTTACCTCCCTCCCTCCAGAGAGCGGAGTTCATGTGAGGGAGGTGCTG 5357
QY 403 -----ArgAlaGlnGlyThrHisValLeuValHisCysLysMetGly 416
Db 5358 AACCTGGCCTGCTTCCACAGACAGGACACCCAGCTCTGTCTCCTCAGATGAGC 5417
QY 417 ValSerArgSerAlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrIleGlySerLeu 436
Db 5418 GTGAGCGGCTCAGCGGCCACAGCTGTGGCTATGTCATGAACAGATGAGAGCCTG 5477
QY 437 GluGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPhe 456
Db 5478 GAGCAGGCGCTCGCCACCTGAGAGAGCTCGCCCATCGCCGCCCAACCTGGCTTC 5537
QY 457 LeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAlaArg 470
Db 5538 CTGCGCCAGCTCAGATCTACCAAGGCACTCTGACGGCCAG 5579
RESULT 13
AB072356 2280 bp mRNA linear PRI 29-JAN-2002
LOCUS AB072356
DEFINITION Homo sapiens mRNA for hssh-1s, complete cds.
ACCESSION AB072356
VERSION AB072356.1 GI:18376660
KEYWORDS
SOURCE
ORGANISM Homo sapiens
cDNA to mRNA.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Nwa, R., Nagata-Ohashi, K., Takeuchi, M., Mizuno, K. and Uemura, T.
Control of actin reorganization by Slingshot, a family of
phosphatases that dephosphorylate Adp/cofilin
Cell 108 (2), 233-246 (2002)
JOURNAL
MEDLINE
21822082
Nwa, R., Nagata-Ohashi, K., Hay, B. A., Takeuchi, M., Mizuno, K. and
Uemura, T.
Direct Submision
TITLE
Submitted (29-SEP-2001) Tadashi Uemura, Laboratory of Molecular
Genetics, The Institute for Virus Research, Kyoto University;
Shogoin-Kawara-cho 53, Sakyo-ku, Kyoto 606-8507, Japan
(E-mail: tuemura@virus.kyoto-u.ac.jp, Tel:81-75-751-4031,
Fax:81-75-751-3989)
FEATURES
source
1..2280
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2280
/gene="hssh-1s"
1..2079
/gene="hssh-1s"
/standard_name="slingshot-1s"
/note="Dual specific phosphatase"
/codon_start=1
/product="hssh-1s"
/protein_id="BAB8415.1"
/db_xref="GI:18376661"
/translation="MALVTIDRSPTPPSAASSASNSLEAGSEEDRKINLSLSEFFM
VKGALFLQGGSSDGGORSLOHPKHADLDLQHLQVMNLRCBDRKLAIVRLSAMA
DVRVMVVVSSGGDTEENILLGVDFSSKSKSTIGMVLRLSDTHDGGGS
VSIAGRMHIFKPVSVQAMMSALQVLRKACEVARHNYPGGVALIMATVYSCSSBO
SCINEMNAMODLESTRPDSPALFVDKPEGRTERTLRKLRSLMSODLENVSKET
RNELEKONCNLKLKEPDIENMLILGOMKPSIIPDHLVLSGEMNASNLEIQQSS
VDYIVNTRLEIDNFPGLFAHNIRVDEETDILLAHMNEATHTNKRHNHSLVH
CKMGSRASSTVIAVYAMKEFGMLPKAVNYVKRSITRPVAGFMROLSEVEGLIDS
KORHNKLWROOTDSLOQPVDDPAGPGGFLPETPGTPEESQLPFLDPAAGPLPPLP
CCFRRLSDPLPSPEDETGSLVHLEDPERALLBEAAPPAVHVRPARPQQGSGLCCK


```
/standard_name="slingshot-1L"
/note="dual specific phosphatase"
/codon_start=1
/product="hSSH-1L"
/db_xref="GI:18376659"
/translation="MALVTIQRSPTPSAASSASNSLEAGSEEDRKLTNLSSEFPM
VKGAALFLQOGSSPOGSIQHPKHAAGDLPQHLQVMI NLRCDRIKILVRLSSAMA
DRVRYAVVYVSSGSDTEENTLLGVDSSEKSRKARRHNTPGVVALIMATYKIDGGGGS
VSTAGRMHIFKPVSVQAMMSALQVHLKACSEKARRHNTPGVVALIMATYKIDGGGGS
SCIENMANMODLESTRPDSPAIFYDKPTBGRTRILKAKRISIMSDOLENTYKSKEL
RNELEKQNMCKLKEFIDENMLLIGQMDKPSLITDHLVISEMNASNLEQSGS
VDILVNTREIDNEPGLFAYHNIRVDEETDILLAHMNEAYHFINAKRMSCLVH
CKMVSASSTVIAVAMKEFGMPLKAYNYVKORSTIRPAGMRQDLESEGTLDAS
KORHNKLMROOTDSLOQVDDPAGPDLPEPTDGPESOLPLDDAOLGPPPLP
CCFRRLSDPLLPSEDETSIVHLEDEPREALIEEAPPAVHAPARPOOGSGTLCCK
DYKRLIEFGSPKSGSSLOVEETEREDEGLAGAWCOLPTDLDNLNENLNNSKR
SCPNEMDDAIFGLINKVPSYKSCADMTPTAGAPASEREDPAPATCTOPAF
LPHITSSPVAILASRSRVPKPSGPTPEPPELPASRRADTSGPAGALPEPASL
LEPSETPKVPLKSLILKNSHCDKNPSTEVILSEVSPKMDPARDRLLEFENSEE
KPTNSYLMQHOESIIQOKAGLVKHTKELERKSYADAPAPSRDGPASRLBASIP
EESODPALHEIPLVMPDSQAGSDESEAPASLEGGSLKPPPEFYRIQDTSXSESP
FKTICTPTSSSSMNLTRSSSDSIHVRKRGVLYKQRIQLETRLRLAGLVYSSP
LKRSHTLAKLSLTSTEDUSSEADPSTVADSDUTTSSESLHEPQGTPRDPATSK
PSCKPAPEMCKSPSMWSK"
```

BASE COUNT 973 a 1100 c 994 g 749 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	3,996-70	Length:	3817
Score:	1113.00	Matches:	232
Percent Similarity:	65.14%	Conservative:	80
Best local Similarity:	48.43%	Mismatches:	127
Query Match:	45.69%	Indels:	40
DB:	9	Gaps:	8

US-09-761-640-4 (1-471) x AB072355 (1-3817)

OY 1 MetalaleuValThValSerArgSerProProGlySerGlyAlaSerThProValGly 20

DB 1 ATGGCCCTGGTGACCCCTGACGCGCTGCCACGCCACGCCGCGCTCCCTGCCGCCAC 60

OY 21 ProTrpAsp-----GlnAlaValGlnArgSerArgLeuGlnArgGln 36

DB 61 AACACGAGTGGAGCTGGCAGCAGCAGAGAATCGAATAATTAACTCAGCTAAGTAG 120

OY 37 SerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56

DB 121 AGCTTTTCATGCGTGAAGGCGCAGCCTCTCTTACCAAGGGA----- 165

OY 57 AlaAlaGluAlaSerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGly 76

DB 166 -----AGCAGCCCTCAA----- 177

OY 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnLysGlnGlu---GlnGln 94

DB 178 -----GCCACAGGAGCTCTCAGCAGCAGCCCAACAACAGCAGGATGATCG 222

OY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeuAla 114

DB 223 CCTCAACATCTTCAGTCAATCAACCTCTCTGCTTCCGAGACAGATCAAGCTGGCA 282

OY 115 AlaGlnLeuGluAlaProArgProProArgLeuArgTyrLeuLeuValVal-----Ser 132

DB 283 GTGCCCTCGAGAGCGCTGGCGGAGCAGCGGCTCCGTCATGCTGTGTGTCACGACGC 342

OY 133 ThrArgGluGlyGlyGlyLeuSerGlnAspGluThrValLeuLeuGlyValAspPhePro 152

DB 343 GGGCGCCAG-----GACACCGAGAGATATCTGCTGGAGTGAGCATTTTCC 390

OY 153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuTrpSerAspThrGln 172

DB 391 AGTAGAAGAAAGTCAACCTGATGATGTTCTCCGACGTGTGAGGACACAGAA 450

OY 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlyGlnSerArgIlePhe 192

DB 451 ATCCACCTTGATGAGGATGCTGCTTACCGCCGACACAGCAGAGATGATGATATT 510

OY 193 LysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGlu 212

DB 511 AAGCTGTGTCTGTCAGCGCATGTGCTGCTGCGCTCAGAGTCTTACAAAGGCTGCCAA 570

OY 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerGlnAlaLeuThrTrpAlaSerHisTyr 232

DB 571 GTGGCCGAGGACACACTACTTCCCGGGGTGAGCTCTATCTGAGCTTACCTAT 630

OY 233 GlnGluArgLeuAspSerGlnGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeu 252

DB 631 GAGACTGATCATGCTCCGAGCAGAGCTGCATCAACAGCTGAGAACCCATGAGACCTG 690

OY 253 GluSerLeuArgPro-----ProSerAlaGluProGlyGlySerSerGluGlnGln 270

DB 691 GAGTCTACGCGCGGCTGCTCCCGCTATTGTGACAAAGCCCACTGAAGGGAAG 750

OY 271 MetGluGlnAlaIleArgAlaGluLeuTrpLysValIleAspValSerAspLeuLys 290

DB 751 ACCGAGCCCTCATCAAGCCAGCTCCGAGACATCATGATGAGCCAGGATCTGAATAAT 810

OY 291 ValThrSerLysGluIleArgGlnAlaLeuGluLeuArgGlyLeuProLeuGlnGln 310

DB 811 GTGACTTCCAAAGACATTCGTATGATTAAGAAACAGATCAATTGATCTTGAAGAA 870

OY 311 TyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330

DB 871 CTCAGAGATTTATGACAAATGAGATGCTACTTATCTTGAGCAGATGAGCAAGCCCTCC 930

OY 331 ArgIlePheProHisLeuTyrIleuGlySerGluTrpAsnAlaIleAsnLeuGluGlu 350

DB 931 CTATCTTCGATCATCTTATCTCGCTGATGAAATGATCAATCTGAGGAGACTG 990

OY 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrPro 370

DB 991 CAGGCTCAGGGGTGATTTACTATTAAATGTTACCAGAAATGATTAATTTTCTCT 1050

OY 371 GluArgPheThrTyrHisAsnValArgLeuTrpAspGluGlnSerAlaGlnLeuLeuPro 390

DB 1051 GCGTATTGTCATATCAATACATCCAGCTGATGAAAGACACAGACTCTCTGCC 1110

OY 391 HisTrpLysGluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeu 410

DB 1111 CACTGGAATGAAGCGTATCATTTTATAAACAAGAGAGAACCATTCCAAGTGCCTG 1170

OY 411 ValHisCysLysMetGlyValSerArgSerAlaIleThrValLeuAlaIleTyrAlaMetLys 430

DB 1171 GTGCTTGCATAAATGCGGTGATGCTGCGCTCCACAGTCATAGCTATGCAATGAG 1230

OY 431 GlnTyrGluCysSerLeuGlnGlnAlaLeuArgHisValGlnGluLeuArgProIleAla 450

DB 1231 GAATTCGGCTGCGCTCGGAAACATATATACATATGTAACAGCAAGCGGAGCTACG 1290

OY 451 ArgProAsnProGlyPheLeuArgGlnGlnIleTyrGlnGlyIleLeuThrAla 469

DB 1291 CGCCCAACGCGGCGCTTATGAGGAGCTGCTGATGATGAAGCATCTTGATGCA 1347

RESULT 15

AX180874

LOCUS AX180874 6374 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 1 from Patent WO0146394.

ACCESSION AX180874

VERSION AX180874.1 GI:15132702

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 6374)

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	2436	100.0	2618	24	AAD36063	Murine dual-specificity phosphatase 15 (DSP-15) cDNA.
2	2436	100.0	2704	24	AB073249	Human MAP kinase p
3	2436	100.0	2704	24	AB073252	Human MAP kinase p
4	2426	99.6	2718	24	AAD36061	Human dual-specificity phosphatase 15 (DSP-15) cDNA.
5	2426	99.6	2781	22	AAAF30485	Human protein phosphatase 15 (DSP-15) cDNA.
6	2426	99.6	2852	24	AB073350	Human MAP kinase p
7	2061.5	84.6	2540	24	AB073351	Human MAP kinase p
8	1987.5	81.6	2332	24	ABN40805	Human MAP kinase p
9	1470.5	60.4	2061	24	ABN59832	Human MAP kinase p
10	1113	45.7	6374	22	AAD09491	Novel human coding
11	1076.5	44.2	2260	22	AAD09493	Human SGP006 phosphatase
12	1045	42.9	1711	22	AAD12966	Human dual-specificity phosphatase 15 (DSP-15) cDNA.
13	1004.5	41.2	1771	22	AAD12965	Human dual-specificity phosphatase 15 (DSP-15) cDNA.
14	980	40.2	1949	22	AAD12965	Human dual-specificity phosphatase 15 (DSP-15) cDNA.
15	918.5	37.7	4467	23	ABL10739	Drosophila melanogaster
16	852	35.0	3468	24	ABLS7466	Human protein phosphatase 15 (DSP-15) cDNA.
17	793	32.6	1026	22	AAAF63578	Human protein phosphatase 15 (DSP-15) cDNA.
18	760.5	31.2	1755	22	AAH14722	Human MAP kinase p
19	760.5	31.2	1755	24	ABL40801	Human MAP kinase p
20	733	30.1	8002	23	ABL10738	Drosophila melanogaster
21	714	29.3	717	22	AAH07057	Human EST-derived
22	689	28.3	599	24	ABL40803	Human MAP kinase p
23	687	28.2	1348	24	ABL40802	Human MAP kinase p
24	648	26.6	1438	22	AAH99712	Human protein phosphatase 15 (DSP-15) cDNA.
25	591.5	24.3	426	24	ABL40800	Human MAP kinase p
26	589	24.2	1052	22	ABL12867	Human MAP kinase p
27	568.5	23.3	571	22	AAH07057	Human EST-derived
28	539	22.1	969	22	AAAF59843	Human novel cytochrome
29	438	18.0	1450	22	AAAF34387	Human novel cytochrome
30	413	17.0	494	24	ABL40804	Human MAP kinase p
31	413	17.0	494	24	ABL40804	Human MAP kinase p
32	401	16.5	447	21	AAAF5065	Human ORFX ORF620
33	357	14.7	951	21	AAAF33441	Human secreted protein
34	269	11.0	828	22	AAAF63567	Murine phosphatase 15 (DSP-15) cDNA.
35	256	10.5	901	22	AAAF63576	Human phosphatase 15 (DSP-15) cDNA.
36	256	10.5	1087	21	AAAF63094	Human cellular protein
37	256	10.5	1292	23	AAAF18101	Human 18221 cDNA.
38	256	10.5	1574	24	AAD2605	Human protein phosphatase 15 (DSP-15) cDNA.
39	235.5	9.7	1337	24	AAAF9409	DNA of APP related
40	235.5	9.7	1830	21	AAAF64262	Human dual-specificity phosphatase 15 (DSP-15) cDNA.
41	235.5	9.7	2192	21	AAAF35762	DNA encoding a human
42	231.5	9.5	687	22	AAAF2608	Murine DSP-3 variant
43	231.5	9.5	1067	22	AAAF63565	Murine phosphatase 15 (DSP-15) cDNA.
44	227.5	9.3	2050	22	AAAF63577	Human phosphatase 15 (DSP-15) cDNA.
45	227.5	9.3	2118	22	AAAF30479	Human protein phosphatase 15 (DSP-15) cDNA.


```

FH Key Location/Qualifiers
FT CDS 35..1450
FT /tag=a
FT /product="Murine DSP-15 protein"
XX
XX WO200224740-A2.
XX
XX 28-MAR-2002.
XX
XX 19-SEP-2001; 2001WO-US29406.
XX
XX 19-SEP-2000; 2000US-233833P.
XX
XX 18-SEP-2001; 2001US-0955732.
XX
XX (CEPT-) CEPTYR INC.
XX
XX Luche RM, Wei B;
XX
XX WPI; 2002-394127/42.
XX
XX P-PSDB; AAE22733.
XX
XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
XX useful for treating e.g. Duchenne muscular dystrophy, cancer,
XX graft-versus-host disease, autoimmune diseases, allergies, metabolic
XX diseases.
XX
XX Claim 56; Fig 4; 91pp; English.
XX
XX The invention relates to a new isolated dual-specificity phosphatase 15
XX (DSP-15) polypeptide which retains the ability to dephosphorylate an
XX activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
XX that dephosphorylate both phosphotyrosine and phosphothreonine/serine
XX residues. DSP-15 polypeptides may be used to identify agents that
XX modulate DSP-15 activity, where such agents may inhibit or enhance signal
XX transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
XX polypeptides, modulating agents, and/or polynucleotides encoding the
XX polypeptides may be used to modulate DSP-15 activity in a patient, and to
XX ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
XX versus-host disease, autoimmune diseases, allergies, metabolic diseases,
XX abnormal cell growth, abnormal cell proliferation and cell cycle
XX abnormalities. DSP-15 alternate form polypeptides are useful in screening
XX assays for modulators of enzyme activity and/or substrate binding. The
XX present sequence is murine DSP-15 cDNA.
XX
XX SQ Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other;
XX
XX Alignment Scores:
XX
XX Pred. No.: 4 51e-195 Length: 2618
XX Score: 2436.00 Matches: 471
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-761-640-4 (1-471) x AAD36063 (1-2618)
XX
XX 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
XX |||||
XX 35 ATGGCCCTGGTCAACATGAGCGCTTGCCTCCGCGGAGCGGCGCCCTCCAGCCCGTGGGG 94
XX |||||
XX 21 ProTTPASpGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
XX |||||
XX 95 CCTGTGGACACAGCGGCTCCAGCGAAGAGCTCCACTCCAGCGAAGCGAGAGCTTTCGGCTG 154
XX |||||
XX 41 LeuArgGlyAlaValLeuGlnLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnVala 60
XX |||||
XX 155 CTCCTGGGGGCTGCTCTGGAGCTGCGAGATGAGGGGACAAATGATATGACAGAGGCC 214
XX |||||
XX 61 SerSerGluProThrGluLysAlaProSerGluGluGluLeuLeuHisGlyAspGlnThrAsp 80
XX |||||
XX 215 AGTTTGACGCAACAGAAAGGCCCGGAGTGAAGAGAGAGCTCCACGGGGACCGACAGAC 274
XX |||||
XX 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100

```

```

DB 275 TTCGGGCAAGGATCCACAGAGTCCCGAGAGAGGAGGACAGGACGACCTGCACCTC 334
DB 101 MetValGlnLeuLeuArgProGlnAspPrlLeuGlnAlaGlnLeuGlnLysPro 120
DB 335 ATGTATACACTCTGTAGGCGGACAGATGACATCCGCTGGCAGCCAGCTGAGGACACC 394
DB 121 ArgProProArgLeuArgArgLeuLeuValValSerThrArgGluGluGluLeuSer 140
DB 395 CGGCTCCCGGCTCCGCTACCTGCTGAGTGTTCACACGAGAGAGAGAGAGTGTGACC 454
DB 141 GlnAspGluThrValLeuLeuGlnGlyValAspPheProAspSerSerProSerCysThr 160
DB 455 CAGATGACAGCGCTCTCTGGGGCGGATTTCCCTGACAGACGACCTCCCGACCTGACAC 514
DB 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValValLeuAspGluLysGly 180
DB 515 CTGGGCTGTGCTTGGCCCTGTGAGTGCACCCAGGTGTACTGTGATGAGACAGGGGGC 574
DB 181 PheSerValThrSerGlyGlnSerArgGluLeuPheLysProIleSerIleGlnThrMet 200
DB 575 TTCAGGCTACGCTGTGTGGGCAAGCCGAGTCTTCAAGCCATCTCCATCCAGACATG 634
DB 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
DB 635 TGGGCAACACTCCAGGTATTGACACAGCATGTGAGGAGCTCTAGCGAGCGGCTTGTA 694
DB 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrglnGluArgLeuAsnSerGluGln 240
DB 695 CCGGTTGGGAGTCCCTTCACTGCGGACCCACTACAGAGAGAGACTGAACTCCAGACAG 754
DB 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
DB 755 ACTGCTCTCAATGACAGCGAGCTTATGGCCGAGCTGAGAGTCTTGGGCTCCAGCGCC 814
DB 261 GluProGlyGlySerSerGluGlnGluGlnMetGluAlaIleArgAlaGluLeuTrp 280
DB 815 GAGCTGTGGGGGCTCCCAACAGAGAGAGATGAGACAGGACATCCGTGACCTGATG 874
DB 281 LysValLeuAspValSerAspLeuGluSerValThrSerGlyGluLeuArgGlnAlaLeu 300
DB 875 AAGGTTTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 934
DB 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrglnArgAspPheIleAsnGlnMetLeu 320
DB 935 GAGCTCGGCTGGGGCTCCCTCCAGCATGACCTGATCATGACCAACAGATGCTG 994
DB 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrglySer 340
DB 995 CTGCTGTGTGGACAGGGGAGCGAGCTCCCGCATCTTCCCGACCTTACCTGGGCTCA 1054
DB 341 GluTrpAsnAlaAlaAsnLeuGlnGluGlnArgGlnArgValThrHisIleLeuAsn 360
DB 1055 GAGTGAACGACCAACACTCGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1114
DB 361 MetAlaArgGluIleAspAsnPheTyrglnArgPheThrTyrglnHisValArgLeu 380
DB 1115 ATGGCCCGGAGATGTACAACTTCTACCTGAGCGCTTACCTACCAACAAATGTGGCCCTC 1174
DB 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIleGlu 400
DB 1175 TGGGATGAGAGTCCGCTCCAGCTGTGCGGCACTGGAAGAGAGAGAGAGAGAGAGAG 1234
DB 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
DB 1235 GCTGCAAGACACAGGAGGCCACCGACGTGCTGCTCACTGCAAGATGGGCTCAGCCGCTTA 1294
DB 421 AlaAlaThrValLeuAlaValAlaMetLysGlnTyrglnCysSerLeuGlnAlaLeu 440
DB 1295 GCGGCAACAGTGTGCTATGCAATGAGAGAGATGAGATGAGAGAGAGAGAGAGAGAG 1354
DB 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460

```

Db 1355 CGCCACGTCGACGAGCTCCGGCCCATCGCCGCCCAACCTGGCTTCGCGCAGCTG 1414
Oy 461 GlnIleYrGlnGlyIleLeuThrAlaArgThr 471
Db 1415 CAGATCTACAGGCAATCTGACGCGCAGAAC 1447
RESULT 2
AB073249
ID AB073249 standard; cDNA; 2704 BP.
XX
AC AB073249;
XX
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:1.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KM MAP kinase; enzyme; chromosome 11; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..93
FT CDS /*tag= a
FT 94..1509
FT /*tag= b
FT /product= "MAP kinase phosphatase splice form 1"
FT 1510..2704
FT 3'UTR /*tag= c
XX
XX MO200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001; 2001MO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
XX (PEKE) PE CORP NY.
XX
PA Wel M, Ketchum KA, Di Francesco V, Beasley EM;
XX
XX WPI: 2002-575237/61.
DR P-PSDB: ABP51653.
XX
XX Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
XX Claim 1; Fig 1A; 85pp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
CC (1). (1) can be used for identifying a modulator of (1) by contacting
CC (1) with an agent and determining if the agent has modulated the
CC function or activity of (1). (1) is useful for identifying an agent that
CC binds to (1), by contacting (1) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (1). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (1) and the polynucleotide
CC sequences encoding (1) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 1 from the present
XX invention.
XX
SQ Sequence 2704 BP: 569 A; 874 C; 794 G; 467 T; 0 other:
Alignment Scores:
Pred. No.: 4,71e-195 Length: 2704
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0
US-09-761-640-4 (1-471) x AB073249 (1-2704)
Oy 1 MetalAlaValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 94 ATGGCCCTGTGTACAGTAGAGCGGTTCGCCCGGCGACGCGCGCTCCACGCCCTGGGG 153
Oy 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
Db 154 CCTGGGACACAGGCGGTCCAGCAGAGGAGTGCACCTCCAGCAAGGCAAGCTTTCGGGTG 213
Oy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlyAla 60
Db 214 CTCCTGGGGCTGTCTCTGGACTGAGATGGAGGGGCAATGATGATGCACAGAGGCC 273
Oy 61 SerSerGluProThrGluAlaProSerGluGluGluGluGluGluGluGluGluGluGlu 80
Db 274 AGTTCTGAGCCACACAGAGAGAGCCCGAGTGAGAGAGCTCCACGGGGACACAGAC 333
Oy 81 PheGlyGlnGlySerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
Db 334 TTCGGCAAGGATCCACAGATCCCGCAAGCAGAGGAGCAGAGCACCCTGCACCTC 393
Oy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuGluAlaPro 120
Db 394 ATGTGACAGCTGCTGAGGCCCGCAGATGACATCCGCTGGCAGCCACTGGAGGCAACC 453
Oy 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGluGluGluSer 140
Db 454 CGGCTCCCGCGCTCGCTACCTGCTGAGTTCATACAGAGAGAGAGAGAGAGTGTGAGC 513
Oy 141 GlnAspGluThrValLeuLeuGlyAlaAspPheProAspSerSerSerProSerCysThr 160
Db 514 CAGGATGAGAGGCTCTCTGGCGTGGATTCCTTGACAGAGCTCCCGACGTCGACC 573
Oy 161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGly 180
Db 574 CTGGGCTGTGTCTTCCCTCTGGAGTGACACCCAGGTGTACTTGTGATGGAGCGGGGCG 633
Oy 181 PheSerValThrSerGlyGlyGlnSerArgIlePheIysProIleSerIleGlnThrMet 200
Db 634 TTCAGCTGAGCTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCATCGACACATG 693
Oy 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
Db 694 TGGGCCACACTCCAGGATTTGACCAAGCATGTGAGGAGCTTAAGGAGGCGCTTGTGA 753
Oy 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
Db 754 CCGGCTGGCAGTGGCTCTGCACTGGGCCAGCCACTACAGAGAGACTGAACCTCGAACAG 813
Oy 241 SerCysLeuAsnGluTyrTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
Db 814 AGCTCCTCAATGAGTAGAGCAGGCTATGGCCGACCTGAGATCTCTCGGCTCCAGGCC 873
Oy 261 GluProGlyGlySerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
Db 874 GAGCTGGCGGGTCTCTGCAACAGAGAGAGATGAGGAGCGGAGTCCGGCTGAGCTGGG 933
Oy 281 LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
Db 934 AAGGTGTGGATGTGACGTGACCTGGAGAGTGTCACTTCCMAAGAGATCCGCCAGGCTGTG 993
Oy 301 GluLeuArgLeuGlyLeuProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
Db 994 GAGCTGGCCCTGGGGCTCCCTCCAGCAGTACCTGACTCATGACAGACAGATGCTG 1053
Oy 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
Db 1054 CTGCTGTGTGGACACAGGAGGAGGAGCTTCGCGCATTTTCCCGACCTGTACCTGGGCTCA 1113
Oy 341 GluTrpAsnAlaAlaAsnLeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAsn 360

```
Db 1114 GAGTGAACGACGCAACCTGGAGAGCTGCAGAGAACGAGGTCACCATTTTAC 1173
Qy 361 MetAlaArgGluLeuAspAspPheTyrProGlnAArgPheThrTyrHisAsnValArgLeu 380
Db 1174 ATGGCCGGGAGATTGACACTTCTACCTGAGCGCTTCACTACCAATGTCGCCTC 1233
Qy 381 TrpAspGluLeuSerAlaGlnLeuLeuProHisTrpLysGluThrHisAsnPheLeu 400
Db 1234 TGGGATGAGAGTGGCCGACGCTGCTCCGACACAGAGAGACGCCGCTTCATTGAG 1293
Qy 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1294 GCTGCAAGACACAGGCGACCCAGCTGCTGTCACAGTGAAGTGGCGCTCAGCCGCTCA 1353
Qy 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrGlnGlySerLeuGlnAlaLeu 440
Db 1354 GCGGCGACACTGCTGGCTTATGCCATGACCACTAGCAATGCAAGCTGGAGCAGGCCCTG 1413
Qy 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1414 CGCCAGCTGAGAGAGCTCCGCGCCATCGCCGCCCAACCTGGCTTCTGGCCAGCTG 1473
Qy 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
Db 1474 CAGATCTACAGGCGCATCTCTGACGGCCAGAAC 1506
RESULT 3
AB073252
ID AB073252 standard; cDNA: 2704 BP.
XX
AC AB073252;
XX
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism;
KW SNP; gene; ss.
XX
OS Homo sapiens.
XX
FH Key 1. 93 Location/Qualifiers
FT 5'UTR /tag= a
FT CDS /tag= b
FT /product= "MAP kinase phosphatase splice form 1"
FT variation /tag= c
FT /replace(577,A)
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT variation /tag= d
FT /replace(1451,G)
FT /tag= e
FT 3'UTR /standard_name= "single nucleotide polymorphism (SNP)"
FT variation /tag= f
FT /replace(2641,A)
FT /tag= f
FT /standard_name= "single nucleotide polymorphism (SNP)"
XX
FN WO200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
XX
PR 18-JAN-2001; 2001US-0761640.
XX
PA (PEKE ) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
```

```
XX
DR MPI: 2002-575237/61.
DR P-PSDB; ABP51653.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
PS Claim 1; Fig 3A; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I) which can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 1 from the present
CC invention.
XX
SQ Sequence 2704 BP: 569 A; 874 C; 794 G; 467 T; 0 other:
XX
Alignment Scores:
Pred. No.: 4,71e-195 Length: 2704
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
XX
US-09-761-640-4 (1-471) x AB073252 (1-2704)
Qy 1 MetaAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 94 ATGGCCCTGTCACAGTACAGCGTTCGCCGCCGCGGACGCGCTCCACGCCCTGGGG 153
Qy 21 ProTrpAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Db 154 CCCTGGACACAGGCGCTCCAGCGAAGAGTGCATCCAGCAGAGCGAGAGCTTGGCGTG 213
Qy 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAsnAspAlaAlaGlnAla 60
Db 214 CTCGGTGGGCTGCTCCTGGGACTGCTGAGATGAGAGGACAAATGATGACGACGAGGCC 273
Qy 61 SerSerGluProThrGlnLysAlaProSerGlnGlnLeuHisGlyAspGlnThrAsp 80
Db 274 AGTCTGAGCCACAGAGAGGCCCGAGTGAAGAGAGCTCCAGCGGAGACAGACAGAC 333
Qy 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeu 100
Db 334 TTGGGGCAGAGATCCAGAGTCCCAAGACAGAGAGGAGGCGACCTGCACCTC 393
Qy 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuGlnAlaPro 120
Db 394 ATGCTACAGCTGCTGAGGCGCAGAGATGACATCCGCTGACACCCAGCTGGAGGACCC 453
Qy 121 ArgProProArgLeuArgTyrLeuLeuValSerThrArgGlnGlyGlnLeuSer 140
Db 454 CGGCTCCCGCGCTCCCTACCTGCTGAGTTCATACAGAGAGAGAGAGAGTCTGAGAG 513
Qy 141 GlnAspGluThrValLeuLeuGlnGlyValAspPheProAspSerSerProSerCysThr 160
Db 514 CAGAGTACAGAGGCTCTCTGGGCGTGGATTCTCTGACAGCAGCTCCCACTGTCACAC 573
Qy 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyGly 180
Db 574 CTGGGCTGCTGGTGGCCCTGCTGAGTGAACCCAGAGTGTACTTACATGAGACGGGGC 633
Qy 181 PheSerValThrSerIleGlyGlnSerArgGlnPheLysProIleSerIleGlnThrMet 200
```

```

Db 634 TTCAGCGTCAGCTCTGTGGGCAAAACCCGATTCACGCCCATCTCCATCCAGCATG 693
OY 201 TTPAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlnGlySerGlyVal 220
Db 694 TGGGCCACACTCCAGTATTGCAACAGCATGTGAGGCGAGCTCTAGGACGGCCTTGTA 753
OY 221 ProGlyGlySerAlaLeuThrTTPAlaSerHisTYrGlnGlnIleArgLeuAsnSerGlnGln 240
Db 754 CCGGGTGGCAGTGGCCCTCACTGGGGCAGCCACTACAGAGAGAGACTGAACCTCCAGAC 813
OY 241 SerCysLeuAsnGlnTTPThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
Db 814 AACTGCTCAATGAGTGGAGGGCTATGGCCGACCTGGAGTCTCTGGGGCTCCAGCGCC 873
OY 261 GluProGlyGlySerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 874 GAGCCTGGGGGTCTCTCAACAACAGGAGATGAGACAGGCGATCCGTCTGAGCTGTGG 933
OY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerGlnIleArgGlnAlaLeu 300
Db 934 AAAGTGTGATGTCAGTACCTGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTGTG 993
OY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnGlnMetLeu 320
Db 994 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGTGAAGCTTATCCACACAGATGCTG 1053
OY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySer 340
Db 1054 CTGCTGGTGGCAGAGGGAGCGAGCTCCGATCTCCGACCTCTCACTGAGGCTCTCA 1113
OY 341 GluTTPAsnAlaAlaAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 1114 GAGTGGACAGCAGCAAACTGGAGGAGCTGCAGAGCAACAGGTTCCACCACATCTGAAAC 1173
OY 361 MetAlaArgGlnIleAspAsnPheTyrProGlnIleArgPheThrTyrHisAsnValArgLeu 380
Db 1174 ATGGCCCGGGAATTGACAACTTACCTCGAGCGCTTACCTACCAACAATGTGGCCCTTC 1233
OY 381 TTPAspGlnGlnSerAlaGlnLeuLeuProHisTTPLeuGlnIleThrHisArgPheIleGlu 400
Db 1234 TGGGATGAGAGATCCGCCAGCTGTGGCCGACCTGGAAGAGACGACGCGCTTCATTGAG 1293
OY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1294 GCTGCAGAGCAGACAGGGCACCACGCTGTGTCACACTGCAGATGGCGCTCAGACCCCTCA 1353
OY 421 AlaAlaThrValLeuAlaThrAlaMetLysGlnTyrGlnLysSerLeuGlnAlaLeu 440
Db 1354 GCGGCCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCTGGAGCAGGGCCCTG 1413
OY 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1414 CCGCAGCTGACAGAGCTCCGGGCCATCGCCGCCCAACCTGTGCTTCTGTGGCCAGCTG 1473
OY 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
Db 1474 CAGATCTACAGGAGCATCTGACGCGCAGACCC 1506

```

RESULT 4
AAD36061
ID AAD36061 standard; cDNA, 2718 BP.

AC AAD36061;

DT 09-AUG-2002 (first entry)

DE Human dual-specificity phosphatase 15 (DSP-15) cDNA.

XX Human dual-specificity phosphatase 15; DSP-15; antiallergic; cytostatic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; chromosome 11q; gene; ss.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 35..2014
XX FT /*tag= a
XX FT /product= "Human DSP-15 protein"
XX PN MO200224740-A2.
XX PD 28-MAR-2002.
XX PF 19-SEP-2001; 2001WO-US29406.
XX PR 19-SEP-2000; 2000US-233833P.
XX PR 18-SEP-2001; 2001US-0955732.
XX PA (CEPT-) CEPTYR INC.
XX PI Luche RM, Wel B;
XX P1 MPI: 2002-394127/42.
XX DR P-PSDB: AAE22729.
XX PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
XX PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
XX PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
XX PT diseases
XX PS Claim 7; Fig 1; 91pp; English.
XX CC The invention relates to a new isolated dual-specificity phosphatase 15
XX CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
XX CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
XX CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
XX CC residues. DSP-15 polypeptides may be used to identify agents that
XX CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
XX CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
XX CC polypeptides, modulating agents, and/or polynucleotides encoding the
XX CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
XX CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
XX CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
XX CC abnormal cell growth, abnormal cell proliferation and cell cycle
XX CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
XX CC assays for modulators of enzyme activity and/or substrate binding. The
XX CC present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
XX CC chromosome 11q.
XX SQ Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;

```

Alignment Scores:

Pred. No.:	3,29e-194	Length:	2718
Score:	2426.00	Matches:	469
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.59%	Indels:	0
DB:	24	Gaps:	0

US-09-761-640-4 (1-471) x AAD36061 (1-2718)

```

OY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
XX |
XX |
XX |
Db 35 ATGGCCCTGTCACAGTGAAGCGTTCGCCCGGCGAGCGGCGCTTCCACGCCCGTGGGG 94
OY 21 ProTTPAspGlnAlaValGlnArgArgSerArgLeuGlnIleArgGlnSerPheAlaVal 40
XX |
XX |
XX |
Db 95 CCTGGGACCGAGCGGTCCAGCGAAGAGTGCACCTCCAGGAGCGAGGCTTGGGGTG 154
OY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
XX |
XX |
XX |
Db 155 CTCCTGGGGGCTCTCTGGGACTGCAGATGAGGGGCAATGATATGTCAGCAGAGGCC 214
OY 61 SerSerGluProThrGlnLysAlaProSerGlnGlnLeuHisGlnLysGlnThrAsp 80

```

```

Db      215 AGTTCTGAGCAACAGAGAGGCCCCGAGTGAAGAGAGCTCCACGGGGACACAGAC 274
      |||
Oy      81  PhgElYgInglYserInserProGlYngIngluIngluInaArgInHIsLeu 100
      |||
Db      275 TTGGGGCAAGAAATCCAGAAATCCCCAGAACAGAGAGAGAGAGACACCTGCACCTC 334
      |||
Oy      101 MetValGlnLeuLeuArpProGlnAspAsp11earGleuAaIaIaInleuGlnAlaPro 120
      |||
Db      335 ATGGTACAGTGGTGGAGAGCGGAGAGATGACATCCGCTGGCAGCCCGCTGGAGGCCACC 394
      |||
Oy      121 ArgProPArArgLeuArGlyrLeuLeuValValSerThrArpGluGluGluYLeuSer 140
      |||
Db      395 CGGCTCTCCCGGCTCCGCTACCTGCTGGTATTCTACAGAGAGAGAGAAAGCTCTGAGC 454
      |||
Oy      141 GlnAspGlnThrValLeuLeuGluYValAspPheProAspSerSerProSerCysThr 160
      |||
Db      455 CAGAGTAGAGACGGTCTCTCCGGGCGTGGATTTCCTGACAGACAGCTCCCGAGCTGCACC 514
      |||
Oy      161 LeuGlyLeuValLeuProLeuTrpSerAspThrGlnValYrLeuAspGlyAspGly 180
      |||
Db      515 CTGGGCTGTGTCTTGGCCCTCTGAGTAGACACCAGGTACTTAATGAGACGGGGCC 574
      |||
Oy      181 PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
      |||
Db      575 TTCAGGTGACGCTGTGTGGGCAAGCCGATCTTCAGGCCATCTCCATCCAGACCATG 634
      |||
Oy      201 TrpAlaThrLeuGlnValLeuHIsGlnAlaCysGlnAlaIaIaInleuYserGlyLeuVal 220
      |||
Db      635 TGGGCAACACCTCAGGATATTGCACCAAGCATGTGAGGACAGCTTAGCGAGGCCCTTGA 694
      |||
Oy      221 ProGlyYrSerAlaLeuThrTrpAlaSerHIsTYrGlnGluArGleuAsnSerGluGln 240
      |||
Db      695 CCGGGTGGCAGTGGCTTCACTGGCCAGCCACTACAGAGAGAGACTGAACCTCCGAACAG 754
      |||
Oy      241 SerCysLeuAsnGlnTrpThrAlaMetAlaAspLeuGlnSerLeuArpProPArAla 260
      |||
Db      755 AGCTGCTCAATAGATGAGAGGCTATGGCCGAGCTGAGCTGTGGGCTCCACAGCC 814
      |||
Oy      261 GluProGlyYrSerSerGluGlnGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
      |||
Db      815 GAGCTGGCGGGTCTCAGAACAGAGAGATGAGACAGGCGATCCCTGCTGAGCTGTGG 874
      |||
Oy      281 LysValLeuAspValSerAspLeuGlnSerValThrSerLysGluIleArgGlnAlaLeu 300
      |||
Db      875 AAAGTGTGGATGTACGTACCTGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTGTG 934
      |||
Oy      301 GluLeuArGleuGlyLeuProLeuGlnGlnTYrArgAspPheIleAspAsnGlnMetLeu 320
      |||
Db      935 GAGCTGGCGCTGGGCTCCCTCCAGCAGTACCGTACTTCATGCAACACAGATGTG 994
      |||
Oy      321 LeuLeuValAlaGlnArGAspArgAlaSerArgIlePheProHIsLeuTrpLysSer 340
      |||
Db      995 CTGCTGTGGTGGACAGGGGACGAGCCGCCGATCTTCCCACTCTACCTGAGGCTCA 1054
      |||
Oy      341 GluTrpAsnAlaAlaAsnLeuGluLeuGlnArGAsnArGValThrHIsIleLeuAsn 360
      |||
Db      1055 GAGTGAACGCGCAAACTGTGAGAGCTGCAAGAGAAACAGGTCAACCCACATCTTGAAC 1114
      |||
Oy      361 MetAlaArGluIleAspAsnPheTYrProGluArGpPheThrTYrHIsAsnAlaArgLeu 380
      |||
Db      1115 ATGGCCCGGGAATTGACAACTTACCTTGAGCGCTTACCTACACATGAGGCCCTTC 1174
      |||
Oy      381 TrpAspGluGluSerAlaGlnLeuLeuProHIsTrpLysGluThrHIsArpPheIleGlu 400
      |||
Db      1175 TGGGATGAGAGATCGGCCAGCTGCGCGCACTGGAAGGAACACACCGCTTCATTTAG 1234
      |||
Oy      401 AlaAlaArGAlaGlnGlyrThrHIsValLeuValHIsCysLysMetGlyValSerArGser 420
      |||
Db      1235 GCTGAAAGAGCAGAGGAGACCCACGTGCTGTGCTCACTGCMAATGAGGCGCTCAGCCGCTCA 1294
      |||
Oy      421 AlaAlaThrValLeuAlaTYrAlaMetLysGlnTYrGluCysSerLeuGluInAlaLeu 440
      |||

```

```

Db      1295 GCGGCCACACAGTGGCTATGCCATGACAGCATTGCACTGAGCCTGGAGACGCCCCTG 1354
      |||
Oy      441 ArgHIsValGlnGluLeuArpProIleAlaArgProAsnProGlyPheLeuArGlnLeu 460
      |||
Db      1355 CGCCACGTCGAGGAGGCTCCGGCCCATGCGCCGCCCAACCTGTGCTTGGCGCACACTG 1414
      |||
Oy      461 GlnIleTYrGlnGlyIleLeuThrAla 469
      |||
Db      1415 CAGATCTACAGGGCATCTGACGGCC 1441
      |||

RESULT 5
AAF30485
ID   AAF30485 standard; cdna; 2781 bp.
XX
AC   AAF30485;
XX
DT   29-MAY-2001 (first entry)
XX
DE   Human protein phosphatase and kinase protein-10 cDNA 5039718CBI.
XX
KW   Protein phosphatase and kinase protein; PPHK-10; human;
KW   gastrointestinal disorder; immune system disorder;
KW   neurological disorder; cell proliferative disorder; cancer;
KW   diagnosis; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   key      location/Qualifiers
FT   CDS      87..2066
FT           /*tag= a
FT           sig_peptide 87..143
FT           /*tag= b
FT           mat_peptide 144..2063
FT           /*tag= c
FT           misc_feature 219..279
FT           /*tag= d
FT           /*note= "unique fragment"
FT           921..980
FT           /*tag= e
FT           /*note= "unique fragment"

WO200120004-A2.
XX
PD   22-MAR-2001.
XX
PF   14-SEP-2000; 2000WO-US25515.
XX
PR   15-SEP-1999; 99US-0154141.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
PI   Lu DAM;
XX
DR   WPI: 2001-244811/25.
DR   P-PSDB: AAB20331.
XX
PT   Novel human protein phosphatase and kinase proteins for diagnosis,
PT   treatment and prevention of gastrointestinal, immune system,
PT   neurological and cell proliferative disorders -
XX
PS   Claim 5; page 101-102; 103pp; English.
XX

The present sequence is that of cDNA encoding novel human
protein phosphatase and kinase protein PPHK-10 (see AAB20331).
The cDNA was initially identified in incyte Clone ID No. 5039718CBI,
from a colon tumor tissue library. Tissues that express PPHK-10
(as a fraction of total tissues expressing PPHK-10) include
reproductive (0.343), gastrointestinal (0.194) and hematopoietic
or immune (0.134). Diseases or conditions associated with tissues
expressing PPHK-10 (as a fraction of total tissues expressing
PPHK-10) include cancer (0.352), inflammation or trauma (0.314) or
cell proliferation (0.090). The encoded protein shows homology to

```



```

XX 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX (PEKE ) PE CORP NY.
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
DR P-PSDB; ABP51654.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
PS Claim 1, Fig 1B; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 2 from the present
CC invention.
XX
SQ Sequence 2852 BP: 674 A; 895 C; 807 G; 476 T; 0 other:
XX
Alignment Scores:
Pred. No.: 3,51e-194 Length: 2852
Score: 2426.00 Matches: 469
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.59% Indels: 0
DB: 24 Gaps: 0
XX
US-09-761-640-4 (1-471) x AB073250 (1-2852)
QY 1 MetalAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
DB 57 ATGGCGCTGTGTACAGTAGAGCGCTTCGCCCGGCGAGCGGCGCTCCAGCGCGGGGG 116
QY 21 ProTPAspGlnAlaValGlnAlaArgSerArgLeuGlnAlaArgGlnSerPheAlaVal 40
DB 117 CCCTGGGACACAGCGGCTCCAGGAGAGTCCAGTCCAGCCAGGAGGAGGCTTGGCGTG 176
QY 41 LeuAlaArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGluAla 60
DB 177 CTCCTGGGGGCTGTCTGGAGCTGAGAGTGGAGGGGAGAAAGATGATGCACCAAGGCC 236
QY 61 SerSerGluProThrGlnLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
DB 237 AGTCTGAGCCAAAGAGAGAGGCCCGGAGTAGAGAGAGCTCCAGGGGAGCCAGACAGC 296
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnAlaArgGlnHisLeuHisLeu 100
DB 297 TTCGGGCAAGGATCCCAAGTCCCAAGAGCAGAGAGCAGAGCAGACACTGCACCTC 356
QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgGluAlaAlaGlnLeuGluAlaPro 120
DB 357 ATGGTACAGCTGCTGAGGCGCAGAGATGACATCCGCTGGAGAGCCCACTGGAGGACCC 416
QY 121 ArgProProAlaArgLeuArgTyrLeuLeuValAlaSerThrArgGluGlyGlyLeuSer 140
DB 417 CGGCTTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
QY 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerProSerCysThr 160
DB 477 CAGGATGAGAGCGTCTCTGCGGTGGATTTCCCTGACACAGCTCCCGAGCTGCACC 536

```

```

QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
DB 537 CTGGCGCTGTGTGCCCCCTGTGAGTGCACCCAGGATGACTTAAAGTGGAGAGGGGGG 596
QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLeuProIleSerIleGlnThrMet 200
DB 597 TTCAGCTGACGCTGTGTGGGCAAGCGGATCTTCAAGCCCAATCCATCCAGACCATG 656
QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
DB 657 TGGGCCACACTCCAGGATATGCACCAAGCATGTGAGGAGAGCTTACGACCGGCTTTGA 716
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
DB 717 CCGGGTGGCAGTCCCTCAGCTGGGCCAGCAGCTACAGAGAGAGTGAATCCGAGACAG 776
QY 241 SerCysLeuAsnGluThrTrpAlaMetAlaAspLeuGlySerLeuArgProProSerAla 260
DB 777 AGCTGCTCAATGAGTGGAGCGGTATGGCCGAGCTGAGTCTGTGGGCTCCAGCGCC 836
QY 261 GluProGlyGlySerSerGluGlnGluMetGluGlnAlaIleArgAlaGluLeuTrp 280
DB 837 GAGCCTGGCGGGTCTCAGAACAGAGCAGATGAGACAGGAGTCCGCTGAGCTGTGG 896
QY 281 LysValLeuAspValSerAspLeuGlySerValThrSerLysGluIleArgGlnAlaLeu 300
DB 897 AAGGTTTGATGTCTAGTACCTGGAGAGTGTCACTTCAAAAGATCCGCGAGGCTCTG 956
QY 301 GluLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
DB 957 GAGCTGCGCTGGGGGCTCCCTCCAGAGTACCGTGAATTCATGATGACACAGATGCTG 1016
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerAlaGlyIlePheProHisLeuTyrLeuGlySer 340
DB 1017 CTGCTGGTGGCACAGGGGAGCGAGCCTCCCACTCTTCCCACTGAGCTGAGCTGCA 1076
QY 341 GluTrpAsnAlaAlaAsnLeuGluGlnLeuGlnAlaArgAsnArgValThrHisIleLeuAsn 360
DB 1077 GAGTGGAAAGCGACAAACCTGGAGAGCTGAGAGAGGAGAACAGGGGTATCCACATCTTGAC 1136
QY 361 MetAlaIaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
DB 1137 ATGGCGCGGAGATTTGACAACTTCTAACCTGAGCCCTTCACTACACAAATGTGGCGCTC 1196
QY 381 TrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIleGlu 400
DB 1197 TGGGATGAGAGTCCGGCCAGCTGCTGCCGCACTGCAAGAGAGACGACCGCTCAATTGAG 1256
QY 401 AlaAlaIaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
DB 1257 GCTGCAGAGACACAGAGGACCCAGCTGCTGTGTCACCTCAAGATGGGGTTCAGCGCTCA 1316
QY 421 AlaIaIaThrValLeuAlaIaTyrAlaMetLysGlnTyrGlyCysSerLeuGlnAlaLeu 440
DB 1317 GCGGCGACAGTGTGGCTTATGCCATGAAAGCATGATGACAGCTGGAGACAGGCCCTG 1376
QY 441 ArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGluLeu 460
DB 1377 CGCCAGCTGCAGAGAGCTCCGGCCCATCCCGCCCGCCCAACCTGGGCTTCTGCGCACCTG 1436
QY 461 GlnIleTyrGlnGlyIleLeuThrAla 469
DB 1437 CAGATCTACAGGCGATCTGACGAGCC 1463

```

RESULT 7

AB073251 ID AB073251 standard; cDNA: 2540 BP.

AC AB073251:

XX 30-SEP-2002 (first entry)

DE Human MAP kinase phosphatase splice form 3 cDNA sequence SEQ ID NO:3.
XX
KW Human: phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..87
FT CDS /*tag= a
FT 88..1314
FT /*tag= b
FT /*product= "MAP kinase phosphatase splice form 3"
FT 3'UTR 1315..2540
FT /*tag= c
XX
PN W0200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001: 2001WO-US42995.
XX
PR 20-NOV-2000: 2000US-0715177.
PR 18-JAN-2001: 2001US-0761640.
XX
PA (PEKE) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI: 2002-575237/61.
DR P-PSDB: ABP51655.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
PS Claim 1; Fig 1C; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (1). (1) can be used for identifying a modulator of (1) by contacting
CC (1) with an agent and determining if the agent has modulated the
CC function or activity of (1). (1) is useful for identifying an agent that
CC binds to (1). By contacting (1) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (1). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) Kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (1) and the polynucleotide
CC sequences encoding (1) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 3 from the present
CC invention.
XX
SQ Sequence 2540 BP; 552 A; 809 C; 739 G; 440 T; 0 other:

Alignment Scores:
Pred. No.: 1,35e-163 Length: 2540
Score: 2061.50 Matches: 408
Percent Similarity: 86.62% Conservative: 0
Best Local Similarity: 86.62% Mismatches: 0
Query Match: 84.63% Indels: 63
Gaps: 1

US-09-761-640-4 (1-471) x ABQ73251 (1-2540)

QY 1 MetalaleuValIthrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
Db ATGGCCCTGGTCCACAGTAGCCGTTCCGCCCGGCGACGCGCTCCACGCCGCTGGGG 147

QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Db CCTGGGACCGAGGGGCTCCACGAGAGTGCATCCAGCAGGACAGACTTTGGGCTG 207

QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGlnAla 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 208 CTCCTGGGGCTGTCTCGGACTGCGAGATGAGGGGACAAATGATGATGACAGAGGCC 267
QY SerSerGluProThrGlnLysAlaProSerGluGluGlnLeuHisGlyAspGlnThrAsp 80
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
268 AGTTCTGAGCCACAGAGAGAGAGCCCGGAGTGCAGAGAGAGCTCCAGGGAGCAGACAGAC 327
QY PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
Db 328 TTCCGGGCAAGAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuAlaPro 120
Db 388 ATGGTACAGCTGCTGAGGCGCGAGATGATACCTCCCTGGGAGCCACAGCTGGAGGACCC 447
QY ArgProProArgLeuArgTyrLeuLeuValIleSerThrArgGluGluGlyLeuSer 140
Db 448 CGGCTCCCGCGCTCGCTACCTGCTGAGTTCATCTTACACGAGAGAGAGAGTCTGAGC 507
QY GlnAspGluThrValLeuLeuGlyAlaAspPheProAspSerSerProSerCysThr 160
Db 508 CAGAGTGAAGAGCGTCTCTCGGCGCTGGATTTCCCTGACAGAGAGCTCCCGCAGCTGCAC 567
QY LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGly 180
Db 568 CTGGGCTGTGCTTCCCTCGCTGAGTGACACCCAGGTGACTTATGATGAGAGAGCGGGGC 627
QY PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
Db 628 TTCAGCGTGAAGTCTGGTGGGCAAGCCGAGTTCAGGCCATTCATCCAGCAGACAG 687
QY TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
Db 688 TGG-----
QY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGln 240
Db 690 -----
QY 241 SerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAla 260
Db 690 -----
QY 261 GluProGlyGlySerSerGlnGlnGluGlnMetGluGlnAlaIleArgAlaGluLeuTrp 280
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
691 -----TCTTCAACAG 738
QY LysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAlaLeu 300
Db 739 AAGGTGTGATGTACAGTACCTGAGAGTGTCACTTCCAAAGAGATCCCGCAGGCTCTG 798
QY GlnLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnGlnMetLeu 320
Db 799 GAGCTGCGCCCTGGGGCTCCCTCCAGAGTACCTGATCATGACACACAGATGCTG 858
QY LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuLysSer 340
Db 859 CTGCTGTGGTGCACAGCGGAGACCGAGCTCCCGCATCTTCCCGCAGCTTACCTGGGCTCA 918
QY GlnTrpAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsn 360
Db 919 GAGTGAAGCGCAGCAACCTGGAGAGCTGCAGAGAGAACAGGGTCAACCATCTTTGAGAC 978
QY MetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeu 380
Db 979 ATGGCCCGGAGATTGAACACTTTCACCTGAGCCCTTACCTACACAAATGTGCGCCTC 1038
QY TrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIleGln 400
Db 1039 TGGGATGAGAGTCCGCCAGCTGCTGCCCACTGGAAGAGAGAGAGAGAGAGAGAGAGAG 1098
QY AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1099 GGTGCAGAGAGACAGAGGAGACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158

QY 421 A1AAlaThrValIleuAlaTyrAlaMetLysGlnTyrGluCysSerLeuGlnIleuAlaLeu 440
 CC |||||||
 DB 1159 GGGGCGACAGTGTGGCTATGGCATGAAGACACTACGAATGCAGCTTGAGCGAGGCCCTG 1218
 QY 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
 CC |||||||
 DB 1219 CCGCAGTGCAGAGAGCTCCGGCCCATCGCCGCCCAACCCCTGCTTCGCGGCCAGCTG 1278
 QY 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
 DB 1279 CAGATCTACCAAGGAGCTCTGACGGCCAGAAC 1311
 RESULT 8
 ABL40805
 ID ABL40805 standard; cDNA; 2322 BP.
 AC ABL40805;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Human MAP kinase phosphatase-like enzyme encoding cDNA.
 XX
 KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
 KW antiasthmatic; antidiabetic; anorectic; cytostatic; cardiant; human;
 KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
 KW neuroleptic; anticonvulsant; anti-HIV; antirhythmic; hypotensive;
 KW antiallergic; dermatological; vulnerability; gene therapy; ss.
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 1..2322
 FT Location/Qualifiers
 FT /tag= a
 FT /product= "MAP kinase phosphatase-like enzyme"
 FT /note= "contains internal codon deletions"
 FT /transl_except= "(pos: 156..157, aa: Leu)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 180..181, aa: Arg)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 414..415, aa: Val)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 1197..1198, aa: Tyr)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 FT /transl_except= "(pos: 2088..2089, aa: Xaa)"
 FT /note= "there is an apparent one codon deletion which
 FT alters the reading frame"
 XX
 PN W0200220732-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 27-AUG-2001; 2001WO-EP09848.
 PF
 PR 07-SEP-2000; 2000US-230709P.
 PR
 XX (FARB) BAYER AG.
 PA
 XX L1ou J;
 XX
 XX WPI: 2002-339802/37.
 DR P-PSDB; ABB07845.
 DR
 PT New human mitogen activated protein kinase phosphatase-like enzyme
 PT polypeptide, regulators of which are useful for preventing, treating
 PT allergies including asthma, diabetes, obesity, cancer and
 PT cardiovascular diseases
 XX
 PS Claim 1; Fig 13; 134pp; English.

XX
 CC The invention relates to a purified human mitogen activated protein (MAP)
 CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
 CC by standard recombinant methodology. The MAP kinase phosphatase-like
 CC enzyme and encoding polynucleotides are useful for screening for
 CC modulators which are used for treating a MAP kinase phosphatase-like
 CC enzyme dysfunction related disease, such as asthma, a central nervous
 CC system disorder, diabetes, obesity, chronic obstructive pulmonary
 CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
 CC to treat allergies including asthma, allergic rhinitis, atopic
 CC dermatitis, and anaphylaxis, central nervous system disorders such as
 CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
 CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
 CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
 CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
 CC including myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
 CC vascular diseases. The enzyme is useful in diagnostic assays for
 CC detecting diseases and abnormalities or susceptibility to diseases or
 CC abnormalities related to the presence of mutations in the encoding
 CC nucleic acid sequences. The present sequence represents the human MAP
 CC kinase phosphatase-like enzyme polypeptide encoding cDNA.
 CC
 SQ Sequence 2322 BP; 467 A; 751 C; 718 G; 386 T; 0 other;
 Alignment Scores:
 Pred. No.: 2e-157 Length: 2322
 Score: 1987.50 Matches: 428
 Percent Similarity: 86.63% Conservative: 6
 Best Local Similarity: 85.43% Mismatches: 27
 Query Match: 81.59% Indels: 42
 DB: 24 Gaps: 8
 US-09-761-640-4 (1-471) x ABL40805 (1-2322)
 QY 4 ValThrValSerArgSerProPro----- 11
 DB 53 GTCACAGTGAGCGGTTCGCCCGCTGGCAGCTGGCGCTCCACCGCCGCTGGGGCCCTG 112
 QY 12 -----GlySerGlyAlaSerThrProValGlyProThrAsp----- 23
 DB 113 GAATTCCTGAGAGAGGGGAGGACAGCCCTCCGCCCTCCACCTCGGGAG-CTGCTCTCT 171
 QY 24 -----GlnAlaValGlnArgSerArgLeuGlnArgGlnSerHealValIleu 41
 DB 172 CGGCAGAGACTGCTGCTCCAGGAGAGACTCGACTCCAGCAGAGA-----GCTTGG 219
 QY 42 Arg-----GlyAlaValLeuGlnLeuGlnAspGlyGlyAspAsnAspAlaAla 58
 DB 220 CGGTGCTCTGTGGGGCTGCTCGGAGCTGCAGAGATGAGAGGGGACATGATGATCAGCA 279
 QY 55 GluAlaSerSerGluProThrGlnLysAlaProSerGlnGlnLeuHis---GlyAsp 77
 DB 280 GAGGCCAGTTCGAGCCACACAGAGAAGGCCCGCAGTGAAGAGACCTCCACGGGGAAAC 339
 QY 78 GlnThrAsp-PheGly-GlnGlySerGlnSerProGlnLysGlnGlnArgGln 97
 DB 340 CAGACAGACTTTCGCTGCGATCCAGAGATCCCCAGAGAAGGAGGAGAGAGCAGC 399
 QY 97 IsoleuHisLeuMetValGlnLeuAlaArgProGln-AspAspIleArgLeuAlaAla--- 115
 DB 400 AACTGCACACC-TCATGAGGCGTGTGAGGCCCGCAGAGATGACATCCGACTTGGAGGCCA 458
 QY 116 GlnLeuGlnAlaProArgProProArg-LeuArgTyr---LeuLeuValValSerThrAr 134
 DB 459 ACCTGAGAGCACCCCGGGCTCCCGGGATCCGATACCTGCTGAGTTCATACAG 518
 QY 134 glnu-glyGlnLysLeuSer-glnAspGlnThrValIleuLeuGly-ValAspPheProAs 153
 DB 519 AGAAGGAGAAAGTGTGAGCCAGAGATGAGACGCTCTCTGGACGTGGATTTCCTGA 578
 QY 153 pSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrPheAspThrGlnVa 173
 |||||||


```
Db 447 ATGTCACAGCTGCTGAGGCCGACGATGACATCCGCTGGCAGCCAGCTGAGACGCC 506
Oy 121 ArgProProArgLeuArgTyrLeuValValSerThrArgIuGlyIuSer 140
Db 507 CGGCTCCCGGCTCGGTACCTGCTGATTTCTACAGAGAGAGAGAGAGAGAGAGAG 566
Oy 141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Db 567 CAGAGTACAGACGGCTCCCTGGGGCTGGATTTCCCTGACAGCAGCAGCTCCCGACCTGACC 626
Oy 161 LeuGlyLeuValLeuProLeuThrSerAspPheGlnValTyrLeuAspGlyAspGly 180
Db 627 CTGGGCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Oy 181 PheSerValThrSerGlyGlnSerArgIlePheLeuProIleSerIleGlnThrMet 200
Db 687 TTCAGCGTACGCTGCTGGTGGGCAAGCCGGATTTCAAGCCATCTCCATCCAGACCATG 746
Oy 201 ThrPalaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuVal 220
Db 747 TGGGGCACACTCCAGTATTTGCACCAAGCATGTGAGGACAGCTCTAGGACGGCTTGTGA 806
Oy 221 ProGlyGlySerAlaLeuThrTrpPalaSerHisTyrGlnIuArgLeuAsnSerGluGln 240
Db 807 CCGGGTGGCAGTGCCTCTACCTGGCCAGCCACTACAGAGAGAGAGAGAGAGAGAGAG 866
Oy 241 SerCysLeuAsnGluThrPheAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
Db 867 ACCTGCTCAATGAGTGGAGGCGGTATGGCCGACTGATCTCTGGGCTCCACAGCC 926
Oy 261 GluProGlyGlySerSerLeuGlnGlnMetGluAlaAlaLeuArgIuLeuThrP 280
Db 927 GAGCTGGGGGCTCTCAACAGGAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
Oy 281 LysValLeuAspVal---SerAspLeuGlnSerValThrSerGlyIuLeuArgIuAla 299
Db 987 AAGGCTTGGAGCTGGAAGACACCTCAAGAGACCATGACATCCAGAGGTCTTCTCTCC 1046
Oy 300 LeuGluLeuArgLeuGlyLeuProLeuGlnGlnTyr 311
Db 1047 CACGAGCTCTTCACATGAGAGAGCTCTGACAGCCCTTC 1082
RESULT 10
AAD09491
ID AAD09491 standard; DNA: 6374 BP.
XX
AC AAD09491;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human SGP006 phosphatase polypeptide encoding DNA.
XX
KW Human; SGP006 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiant;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cytoskeletal;
KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnerary; tranquiliser; antisthmatic;
KW hypertensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; mtgraine; chromosome 12q21.3-q22; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 34..3183
FT /*tag= a
FT /product= "Human SGP006 phosphatase polypeptide"
PN MO200146394-A2.
```

```
XX
XX 28-JUN-2001.
PD
XX
XX 21-DEC-2000; 2000MO-US34736.
PF
XX
XX 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JUN-2000; 2000US-0179301.
XX
XX (SUGEN-) SUGEN INC.
PA
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarshanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI: 2001-418058/44.
DR P-PSDB; AAE04833.
XX
XX Novel phosphatase polypeptide useful for treating cancers,
XX immune-related diseases and disorders, cardiovascular disease, brain or
XX neuronal-associated diseases and metabolic disorders
XX
XX Claim 29; Fig 1; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
XX sequences encoding them, as well as various products and methods useful
XX for the diagnosis and treatment of various phosphatase-related diseases
XX and conditions. Substance that modulates the activity of phosphatase
XX polypeptide is used to treat immune-related diseases and disorders,
XX cardiovascular disease, brain or neuronal-associated diseases and
XX metabolic disorders, including cancers of tissues, cancers of
XX haematopoietic origin, diseases of central and peripheral nervous
XX system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
XX amyotrophic lateral sclerosis, viral infections, infections caused by
XX prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
XX sexual dysfunction, mood disorders, attention disorders, cognition
XX disorders, hypotension, hypertension, psychotic disorders, neurological
XX disorders, dyskinesias and organ transplant rejection. The present
XX sequence is a DNA encoding human SGP006 phosphatase polypeptide. This
XX sequence is classified as dual specificity phosphatase (DSP) and MAP
XX kinase phosphatase (MKP). SGP006 gene maps to chromosomal position
XX 12q21.3-q22.
XX
XX Sequence 6374 BP; 1604 A; 1674 C; 1659 G; 1437 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2.65e-83 Length: 6374
XX Score: 1113.00 Matches: 232
XX Percent Similarity: 65.14% Conservative: 80
XX Best Local Similarity: 48.43% Mismatches: 127
XX Query Match: 45.69% Indels: 40
XX DB: 22 Gaps: 8
XX
XX US-09-761-640-4 (1-471) x AAD09491 (1-6374)
Oy 1 MetalaleuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
Db 34 AAGGCTTGGAGCTGGAAGACACCTCAAGAGACCATGACATCCAGAGGTCTTCTCTCC 93
Oy 21 ProThrAsp-----GlnAlaValGlnArgArgSerArgLeuArgGln 36
Db 94 AACACGAGCTTGGAGCTGCGGACGACGAGAGAGATGCAAAATTAAACCTGAGCTTAGTGAG 153
Oy 37 SerPheAlaValLeuArgIuAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56
Db 154 AGCTTTTCATGATGGAAGAGCGACCCCTCTTCTTACCAACAGGGA----- 198
Oy 57 AlaAlaGlnAlaSerSerGluProThrGlnIuysAlaProSerGlnGluIuLeuHisGly 76
Db 199 -----AGCAGGCCCTCA----- 210
Oy 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnIuysGlnGln---GlnGln 94
```

```

Db 211 -----GGCCAGCGAGTCTTCAGACCCCCACAGCATCAGTGTATCG 255
Oy 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspSerIleArgLeuAla 114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CCTCAACATCTTCAGGTGATGATCAACCTTCTGCGTTCGGAAGACAGATCAAGCTGGGA 315
Oy 115 AlaGlnHisLeuAlaProArgProProArgLeuArgTyrLeuLeuValAla-----Ser 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 GTGCCCTGTGAGAGCGCGTGGCGGACCGCGGTCCGGTACATGTTGGTGTACAGCAC 375
Oy 133 ThrArgGlnGlyGlnGlyLeuSerGlnAspGlnThrValLeuLeuGlyValAspPhePro 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 GGCAGCCACAG-----GACACCGGAGGAATATCTCTGGAGTGCATCTTTCC 423
Oy 153 AspSerSerSerProSerCysThrLeuGlnGlyLeuValLeuProLeuThrSerAspThrGln 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 ACTAAGAAAGTAAAGCTGACACCATGGGATGTTCTCCAGCTGGGAGCACACGAAA 483
Oy 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 ATCCACCTGATGAGAGTGTGGTTCAGCGTTCAGACACAGCAAGAAAGATGCACATATT 543
Oy 193 LysProIleSerIleGlnThrMetThrPalaThrLeuGlnValLeuHisGlnAlaCysGln 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 AAGCTGTCTGTCTCAGGACCATGTGTGCTGCCCTGCAGGTCTTACACAAAGCTGCGAA 603
Oy 213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrThrPalaSerHisTyr 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 GTGGCCCGGAGGCACACTACTTCCCGGGGTGTAGCTCTCATCTGGGCTACCTACTAT 663
Oy 233 GlnGlnTyrLeuAsnSerGlnGlnSerCysLeuAsnGlnTyrThrAlaMetAlaAspLeu 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 GAGACTGATCATGCTCCGACAGCTGCATCAAGATGAGAACGCGCATGACGAGACCTG 723
Oy 253 GluSerLeuArgPro-----ProSerAlaGlnTyrGlyGlySerSerGlnGlnGln 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 GAGTCTACGGCGCCGACTCCCGCGCTATTGTGACAAACCCCACTGAAGCGAAAG 783
Oy 271 MetGlnGlnAlaIleArgAlaGlnLeuTyrPlyValLeuAspValSerAspLeuGlnSer 290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 ACCGAGCGCTCATCAAGCCAGCTCCGAGACATCATATGAGCCAGAGATTAGAAAAT 843
Oy 291 ValThrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeuProLeuGlnGln 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 844 GTGACTTCCAAAGAGATTCGTATGATTAAGAAACAGATTAATTGAACTTGAAGAA 903
Oy 311 TyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 904 CTCAGGAAATTATAGACAAATGAGATGCTACTTATCTTGGGACAGATGACAAAGCCCTCC 963
Oy 331 ArgIlePheProHisLeuTyrLeuGlySerGlnTyrPalaIleAlaAsnLeuGlnGln 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 964 CTTATCTTCATCATCTTTATCTCGCTCTGATGAGATGCAATCCATCTGAGGAGACGTG 1023
Oy 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGlnIleAspAsnPheTyrPro 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1024 CAGGCTCGGGGTGATTAATCTTAAATCTTACAGAGAAATGATATTTTCTCT 1083
Oy 371 GlnArgPheThrThrHisAsnValArgLeuTyrPalaSerGlnLeuSerAlaGlnLeuPro 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1084 GCGTTATTTGCAATATCAACATCCAGTCTGATGAAGAACCAACAGACCTCTCCGCC 1143
Oy 391 HisThrPlyGlnThrHisArgPheIleGlnAlaIleArgAlaGlnGlyThrHisValLeu 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1144 CACTGGAATGAAGCGTATCATTTTATAACAAAGGAGAGAACCATTCACAAAGTCCG 1203
Oy 411 ValHisCysLysMetGlyValSerArgSerAlaIleThrValLeuAlaTyrAlaMetLys 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1204 GTGCAATTCGAAATGAGCGTGCAGTCCGCTGCACAGCATATGAGCTAATGCAATGAAG 1263
Oy 431 GlnTyrGlnCysSerLeuGlnAlaLeuArgHisValGlnGlnLeuArgProIleAla 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1264 GAATTCGGCTGCGCTCTGAAAAGCATATTAATATGTAAGCAAGAGCGACATCAAG 1323

```

```

Oy 451 ArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAla 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1324 CGCCCAACGCGCGCTTTATGAGCAGCTGTCTGAGTATGAAGCATTTGGATGCA 1380
RESULT 11
AAD09493
ID AAD09493 standard; DNA; 2260 bp.
XX
AC AAD09493;
XX
DE 10-SEP-2001 (first entry)
XX
DE Human SGP001 phosphatase polypeptide encoding DNA.
XX
KW Human; SGP001 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; organ disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cytosolic;
KW neuroprotective; antiviral; neurotrophic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; valnerary; tranquilliser; antiaesthetic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome Xp11.1-11.3; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 709..2205
FT FT /*tag= a
FT FT /product= "Human SGP001 phosphatase polypeptide"
XX
PD WO200146394-A2.
XX
PN 28-JUN-2001.
XX
PD 21-DEC-2000; 2000MO-US34736.
XX
PE 21-DEC-1999; 99US-0173255.
XX
PR 28-DEC-1999; 99US-0175766.
XX
PR 25-JAN-2000; 2000US-0178078.
XX
PR 31-JAN-2000; 2000US-0179301.
XX
PA (SUGC-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI. 2001-418058/44.
XX
DR P-PSDB: AAE04835.
XX
PT Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
XX
XX Claim 29; Fig 1; 186pp; English.
XX
CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological

```


PF 01-FEB-2001; 2001WO-US03429.
XX
PR 02-FEB-2000; 2000US-0179886.
XX
PA (CEPT-) CEPTIR INC.
XX
PI Lucie RM, Wei B;
XX
DR MPI: 2001-48887/53.
XX P-PSDB: AAE06775.
PT New isolated dual-specificity phosphatase polypeptide for treating
PT cancer, graft-versus-host disease, autoimmune diseases, allergies,
PT metabolic diseases, abnormal cell growth and abnormal cell
PT proliferation -
XX
PS Claim 56; Fig 3; 81pp; English.
XX
CC The present sequence is human dual-specificity phosphatase (DSP)-13
CC cDNA. Inactivation of mitogen-activated protein kinase (MAP-kinase)
CC is mediated by dephosphorylation at a dual phosphorylation motif by DSP
CC which is referred to as MAP-kinase phosphatase. An agent that modulates
CC DSP is useful for treating a disorder selected from Duchenne muscular
CC dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune diseases,
CC allergies, metabolic diseases, abnormal cell growth, abnormal cell
CC proliferation and cell cycle abnormalities. DSP is useful for
CC identifying antibodies and other agents that inhibit DSP-12 and/or
CC DSP-13 activity. DSP and the agents identified are useful for
CC modulating cell proliferation, differentiation and survival. DSP is
CC useful in screening assays for modulators of enzyme activity and
CC substrate binding and for dephosphorylating a substrate of DSP-12 or
CC DSP-13.
XX
SQ Sequence 1711 BP; 514 A; 396 C; 390 G; 411 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,42e-78 Length: 1711
Score: 1045.00 Matches: 200
Percent Similarity: 66.37% Conservative: 94
Best Local Similarity: 45.15% Mismatches: 107
Query Match: 42.90% Indels: 42
DB: Gaps: 4
US-09-761-640-4 (1-471) x AAD12966 (1-1711)
QY 36 GlnSerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyAlaAspAsnSP 55
DB 240 GAGAGCTTTCTAACTGTCAAAGGTGCTCCCTTTTCTACACAGG----- 284
QY 56 AspAlaAlaGluAlaSerSerGlnProThrGlnLysAlaProSerGlnGlnLueuHis 75
DB 284 ----- 284
QY 76 GlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlnLueuHis 94
DB 285 -----GGAATGCGTCATCCACCAAGCAATGAGCCAGAGAGGAAAC 326
QY 95 -----ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGln 108
DB 327 AAGCATGACGAGGATCTCCACACAGCATCTCCAAAGCAATGCTATTCTCCGCCAGAA 386
QY 109 AspAspIleArgLeuAlaGlnLeuGlnLueuAlaProArgProProGlnLeuArgTyrLeu 128
DB 387 GACACACATCAGCGTGGCGTAAGACTGGAAGTACTTACCAAGAAATCGAACACGCTATATG 446
QY 129 LeuValValSerThrArgGlnGlyGlnGlyLeuSerGlnAspGlnThrValLeuLeuGly 148
DB 447 GTAGTGTTTCAACTAATGAGTACAA-----GACACTGGAAGAAGCATCTCTAGGA 500
QY 149 ValAspPheProSerSerSerProSerCysThrLeuGlnGlyLeuValLeuProLeuTyr 168
DB 501 ATGGATTCTCTCTTAATGACAGTACGACTGTACATGAGGCTTAGTTTGCTCTCTGG 560

QY 169 SerAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGln 188
DB 561 AGCGACACGGCTAATTCATTGATGGATGGTGGTTCAGTGATGACGAGATGAACAGA 620
QY 189 SerArgIlePheLysProIleSerIleGlnThrMetThrPalaThrLeuGlnValLeuHis 208
DB 621 GTTCACATATTCMAACCTGATCTGTGCGAGCAATGTGTGTGCTACAGAGCTTACAC 680
QY 209 GlnAlaGlySGluAlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrp 228
DB 681 AAGCTTTGAGAGTGGCCAGAGCCATACCTACTACCCAGGACGCTATTCTTCACTTGG 740
QY 229 AlaSerHisTyrGlnGlnLueuArgLeuAsnSerGlnSerCysLeuAsnGlnTyrThrAla 248
DB 741 GTGAGTTTATTAAGAGCCATATCAACTCAGATCAATCATCTCTGCAATGATGAATGCA 800
QY 249 MetAlaAspLeuGlnSerLeuArgPro-----ProSerAlaGlnProGlyLysSer 266
DB 801 ATGCAAGATGTACAGTCCACCGCCGACCTCTCCAGCTCTTCAACGACATACCTACT 860
QY 267 GlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLueuTyrLysValLeuAspValSer 286
DB 861 CACGTGACACGACAGAAAGCTAATTAAACCAATTAAGGAGATCTGATGCAAGAG 920
QY 287 AspLeuGlnSerValThrSerLysGlnLueuArgGlnAlaLeuGlnLueuArgLeuGlyLeu 306
DB 921 GATTGGAGAAATATTACATCCAAAGATAGACAGATGGAATGGAATGGTGTC 980
QY 307 ProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArg 326
DB 981 AACTTGGCGGAATTCAGAGAAATTTATAGACATGAATGATGATCTCTGTGTAATG 1040
QY 327 AspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGlnTyrPheAsnAlaIleAsn 346
DB 1041 GATAGCCCTACACAGATATTGATGATGCTGTCTGGGCTCGAATGATGATGCTTCACAA 1100
QY 347 LeuGlnGlnLueuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgLueuLeuAsp 366
DB 1101 TTAGAGAGACTTACAGAACCGAGGAGTATATCTTAATGATCACCCTCGAGAGATAGAT 1160
QY 367 AsnPheTyrProGlnLueuArgPheThrThrHisAsnValArgLeuTyrAspGlnLueuSerAla 386
DB 1161 AACTTCTTCCCGAGGAGTCTTGTAGTATCATTAACATTCGCGTATGATGTAAGACGCAACG 1220
QY 387 GlnLeuLeuProHisTyrLysGlnThrHisArgPheIleGlnAlaIleArgAlaGlnGly 406
DB 1221 GATCTCTGCGCTTACGATGACACTTACAAATTCATCTTAAAGCAAGAAACATGGA 1280
QY 407 ThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaAlaThrValLeuAla 426
DB 1281 TCTAATGCTTGTGACACTGCAAAATGGGGGTGAGTCCGCTCACGCTCCACCGTGAATGCC 1340
QY 427 TyrAlaMetLysGlnTyrGlnCysSerLeuGlnAlaLeuArgHisValGlnLueu 446
DB 1341 TATGCAATGAAGAAATATGCTGTAATCTGACACGACCTATGATGATGAAAGAAAGA 1400
QY 447 ArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIle 466
DB 1401 CGAAGGTAACCAAGCCCAACCAAGCTTATGACACATGGAAGATGAGAGGGGATC 1460
QY 467 LeuThrAla 469
DB 1461 TTGCTGGCA 1469
RESULT 13
AAH22966
ID AAH22966 standard; cDNA; 1771 BP.
AC AAH22966;
XX
XX
DT 17-SEP-2001 (first entry)
XX Human phosphatase (PP) encoding cDNA (clone ID 637331CB1).

XX Phosphatase; PP; Alzheimer's disease; Huntington's disease; human;
 KW immune disorder; cancer; neurotropic; immunomodulatory; cytostatic;
 KW neuroprotective; anticonvulsant; ss.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 24..1478
 XX FT /*tag= a
 XX FT /product= "phosphatase"
 FT
 XX MO200153469-A2.
 XX
 XX 26-JUL-2001.
 XX
 XX 18-JAN-2001; 2001WO-US02088.
 XX
 XX 21-JAN-2000; 2000US-0177719.
 XX 28-JAN-2000; 2000US-0178988.
 XX 25-FEB-2000; 2000US-0184959.
 XX 17-MAR-2000; 2000US-0190142.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Bandman O, Tang YT, Azimzai Y, Yue H, Baughn MR, Hillman JL,
 PI Lal P, Wang E, Gandhi AR, Policky JL, Mathur P;
 DR WPI: 2001-451905/48.
 DR P-PSDB: AAB83563.
 XX
 XX Novel phosphatases useful for treating diseases associated with
 PT decreased expression of functional phosphatases, e.g., Alzheimer's
 PT disease, Huntington's disease, immune disorders, and cancers -
 XX
 XX Claim 5: Page 102-103; 103pp; English.
 PS
 CC The invention provides human phosphatases (PP) and polynucleotides
 CC encoding the phosphatases. The polypeptides can be expressed by standard
 CC recombinant methodology. The PP are useful for treating a disease or
 CC condition associated with decreased expression of functional
 CC phosphatases. Compositions containing agonists or antagonists of PP may
 CC be used to treat a disease associated with decreased expression or
 CC overexpression of PP, respectively. Such diseases may include
 CC Alzheimer's disease, Huntington's disease, immune disorders, and
 CC cancers. The present sequence represents a human phosphatase encoding
 CC CDNA.
 XX
 XX Sequence 1771 BP; 529 A; 395 C; 378 G; 469 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 6.44e-75 Length: 1771
 Score: 1004.50 Matches: 199
 Percent Similarity: 64.04% Conservative: 93
 Best Local Similarity: 43.64% Mismatches: 119
 Query Match: 41.24% Indels: 45
 DB: 22 Gaps: 5
 US-09-761-640-4 (1-471) x AAH22966 (1-1771)
 QY 1 MetaLalaLeuValThyValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
 DB 24 AAGGCTTTGGTACGCGTCCAGCGGTCACTACCCGACGACGACCTCCGCGCC 83
 QY 21 ProTPASPGLAAlaValGlnArgArgSerArgLeuGlnArgArg-----GlnSer 37
 DB 84 TCGGAGCGACAGACGTGGGAGGAATGCGGTCAAGCCGACGAGCATCAGCGAGAGC 143
 QY 38 PheAlaValLeuArgGlyAlaValLeuGlnArgGlnArgGlnArgGlnArgGlnArgGln 57
 DB 144 TTTCCTAAGCTGCAAGAGTCTGCTTTTCTTCCACAGC-----182
 QY 58 AlaGlnAlaSerSerGluProThrGlnLysAlaProSerGlnGlnGlnLysGlnGlnGln 77

DB 182 -----182
 QY 78 GlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlnGln-----94
 DB 183 -----GGAANTGGCTATCCACCCAGAGATCAGCCACAGAGCAACAGCAT 230
 QY 95 -----ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAsp 110
 DB 231 GCAGGCGATCTCCACAGCATCTCCACAGCATGTTCTATCTACCTCGCCGACAGAACAC 290
 QY 111 IleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArgProArgGlnArgGlnVal 130
 DB 291 ATCAGCGCTGGCTGTAAGAGTGAAGTACTTACAGAAATGCAACAGCATATGAGTAGTG 350
 QY 131 ValSerThrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 150
 DB 351 GTTTCACACTAATGATGACAA-----GACACTGAAGAAGACATGCTCTAGGATGAT 404
 QY 151 PheProAspSerSerSerProSerCysThrLeuGlnLeuValLeuProLeuThrSerAsp 170
 DB 405 TTCTCCTTAATGACAGTAGACACTTGTACATGGCGCTTAGTTGGCTCTCGAGCGAC 464
 QY 171 ThrGlnValTyrLeuAspGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 190
 DB 465 ACGCTAATTCATTTGGTGGTGTGATGGTTCAGTGTGATGACAGATACAGAGTTTAC 524
 QY 191 IlePheLysProIleSerIleGlnThrMetTyrPalaThrLeuGlnValLeuHisGlnAla 210
 DB 525 ATATTCAAACTGTATCTGTGACAGCATGTGTGTCTGCTACAGAGCTTACACAAAGCT 584
 QY 211 CysGlnAlaAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 230
 DB 585 TGTGAAGTCGCCAGACGCGATACACTACCCAGGAGCGCTATTCTCTCGGGTAGT 644
 QY 231 HisTyrGlnGlnArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 250
 DB 645 TATTATGAGAGCCATATCACTGATCAATCTCAGTCAATGATGATGATCAATGCA 704
 QY 251 AspLeuGlnSerLeuArgPro-----ProSerAlaGlnProGlnGlnGlnGlnGln 268
 DB 705 GATGTACATCTCCACGCGCGCCGACCTCTCCAGCTCTTACCGAGACTACTACTGACCT 764
 QY 269 GlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTyrPheValLeuAspValSerAspLeu 288
 DB 765 GAACGACAGAAAGGCTATTAACCAATTAAGGAGATGATGATGACGAGAAATTGG 824
 QY 289 GlnSerValThrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeuGlnGlnGln 308
 DB 825 GAGAAATATTACATCCAAAGAGATAGACAGAGTTGAAATGCAATGGTGTGCAACTTG 884
 QY 309 GlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArgAspArg 328
 DB 885 CGGGAATTCAAGAAATTTATGACATGAATGATGATGATCTGCTGCAATGGATGAC 944
 QY 329 AlaSerArgIlePheProHisLeuTyrLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 348
 DB 945 CCTACACAGATATTTGAGCATGTCTCGGCTCGAGATGGAATGCAATCTTCAATTAAG 1004
 QY 349 GlnLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGlnIleAspAsn 368
 DB 1005 GACTTACAGAACCGAGGAGGTATTTGTAATGTCTGACCTGACAGATAGATTAATCTTC 1064
 QY 369 TyrProGlnArgPheThrTyrHisAsnValArgLeuTyrPaspGlnGlnGlnGlnGln 388
 DB 1065 TTCCCAAGAGCTTTTACAGATCATTAACATTCGGGTATGATGATGAAGCAACGATCTC 1124
 QY 389 LeuProHisTyrLysGlnThrHisArgPheIleGlnAlaAlaArgAlaGlnGlnGln 408
 DB 1125 CTGCGCTACTGCAATGACATCTCAATTCATCTTAAGCAAAAGATGATGATTA 1184
 QY 409 ValLeuValHisCysLysMetGlnValSerArgSerAlaAlaThrAlaLeuAlaTyrAla 428

Dd	1185	TGCTTGTGCACATCGCAAAATGGGGGCGTAGTCGCTCAGCCGCCACCGCTGAATTGGCATGCA	1244
Oy	429	MetylscintyltyrGlucysSerleuglucInAlaleNarghisValcIn	444
		:	:
		:	:
		:	:
Dd	1245	ATGAGAAGATATGTGCTGGAATCTGGACCGAGCCTCGACCGAATTCGG	1292
RESULT 14			
AAD12965			
ID	AAD12965 standard; cDNA; 1949 BP.		
XX			
AC	AAD12965;		
XX			
DT	16-OCT-2001 (first entry)		
DE			
XX			
KW	Human dual-specificity phosphatase (DSP)-12 cDNA.		
KW	Human: dual-specificity phosphatase; DSP-12: dual phosphorylation motif;		
KW	mitogen-activated protein kinase; MAP-kinase; graft-versus-host disease;		
KW	GvHD; Duchenne muscular dystrophy; therapy; cancer; autoimmune disease;		
KW	allergy; metabolic disease; cell growth; cell proliferation; cytostatic;		
KW	cell cycle abnormality; cell differentiation; anti-allergic; muscular;		
KX	immunosuppressive; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	95..1753	
FT		/tag- a	
FT		/product= "Human dual-specificity phosphatase (DSP)-12"	
XX			
PN	WO200157221-A2.		
XX			
PD	09-AUG-2001.		
XX			
PE	01-FEB-2001; 2001MO-USO3429.		
XX			
PR	02-FEB-2000; 2000US-0179886.		
XX			
PA	(CEPT-) CEPTYR INC.		
XX			
PI	Luche RM, Wei B;		
XX			
DR	WP1: 2001-488887/53.		
XX			
P-PSDB:	AAE06772.		
XX			
PT	New isolated dual-specificity phosphatase polypeptide for treating		
PT	cancer, graft-versus-host disease, autoimmune diseases, allergies,		
PT	metabolic diseases, abnormal cell growth and abnormal cell		
PT	proliferation -		
XX			
PS	Claim 7; Fig 1; 81pp; English.		
XX			
CC	The present sequence is human dual-specificity phosphatase (DSP)-12		
CC	cDNA. Inactivation of mitogen-activated protein kinase (MAP-kinase)		
CC	is mediated by dephosphorylation at a dual phosphorylation motif by DSP		
CC	which is referred to as MAP-kinase phosphatase. An agent that modulates		
CC	DSP is useful for treating a disorder selected from Duchenne muscular		
CC	dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune disease		
CC	allergies, metabolic diseases, abnormal cell growth, abnormal cell		
CC	proliferation and cell cycle abnormalities. DSP is useful for		
CC	identifying antibodies and other agents that inhibit DSP-12 and/or		
CC	DSP-13 activity. DSP and the agents identified are useful for		
CC	modulating cell proliferation, differentiation and survival. DSP is		
CC	useful in screening assays for modulators of enzyme activity and		
CC	substrate binding and for dephosphorylating a substrate of DSP-12 or		
CC	DSP-13.		
XX			
SQ	Sequence 1949 BP; 516 A; 509 C; 513 G; 411 T; 0 other;		
Alignment Scores:			
Pred. No.:	8.4e-73	Length:	1949
Score:	980.00	Matches:	183
Percent Similarity:	73.94%	Conservative:	61

Best Local Similarity:	55.45%	Mismatches:	84
Query Match:	40.23%	Indels:	2
DB:	22	Gaps:	1
US-09-761-640-4 (1-471) x AAD12965 (1-1949)			
QY	142 ASPGIUTHTHValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThrLeu	161	
DB	32 GAGGAGAAATATCTTCGCTGGAGGTGACTTTCACAGTAAGAAAGTAAAGCTGCACCAT	91	
QY	162 GlyLeuValLeuProLeuThrSerAspPheThrLlnValTyrLeuAspGlyAspGlyPhe	181	
DB	92 GGGATGGTTCTCCGACTGTGAGGACACGAAATTCACCTTGTAGGAGATGGGTTC	151	
QY	182 SerValThrSerGlyGlyGlnSerArqIlePheLeuProIleSerIleGlnThrMetTrp	201	
DB	152 AGCGAGACACACGAGGAAGATGACATATTTAAGCTGTCTGTCCAGGCCATGTGG	211	
QY	202 AlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaLeuGlySerGlyLeuValPro	221	
DB	212 TCTGCCCTGCGAGGTCCTTACAGAGCCCTGCCAAGTGGCCCGAGGACCACTACTCTCCC	271	
QY	222 GlyGlySerAlaLeuThrThrAlaSerHisTyrGlnGlnAlaTyrLeuAsnSerGlnGlnSer	241	
DB	272 GGGGTGTAGCTCTCATCTGTGGACTTACCTATGATAGAGACTGCATCATGCTCCGACCAAGC	331	
QY	242 CysLeuAsnGlnTyrThrAlaMetAlaAspLeuGlnSerLeuAsnPro-----ProSer	255	
DB	332 TGCATCAACGAGTGAACGACCCATGAGAGCTGTGAGTACGCGGCCGACTCCCGCGG	391	
QY	260 AlaGluProGlyGlySerSerGlnGlnGlnMetGlnAlaIleArgAlaGluLeu	279	
DB	392 CTATTGTGGACAACCCACCTGGAAGGGAAGAGACCGAGCCCTCATCAAGCCCAAGCTC	451	
QY	280 TrpIysValLeuAspValSerAspLeuGlnSerValThrSerGlyIleArgGlnAla	299	
DB	452 CGAACCATCATGTATGAGCCAGAGATTTTCAAAAATGTGACTTCCAAAAGATTCGTAATGAA	511	
QY	300 LeuGlnLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMet	319	
DB	512 TTAGGAACACAGATAAATTGTACTTGAAAGACATCAAGAAATTTATAGACAAATGACATG	571	
QY	320 LeuLeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGly	335	
DB	572 CTACTTATCTTGGGACAGAGATGAGACAGCCCTCCCTATCTTCGATCATCTTATCTCGGC	631	
QY	340 SerGluTyrAsnAlaAlaAsnLeuGlnGluLeuGlnArgAsnArgValThrHisIleLeu	355	
DB	632 TCTGATGGAAATGCTCCAACTGTGAGGAACCTGCAAGGCGTAGGGGTGATTACATTTTA	691	
QY	360 AsnMetAlaArgGlnIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArg	379	
DB	692 NATGTGTACCAGGAATGATATTTTCTTCTGGCTTATTTGATTCATATACATCCGA	751	
QY	380 LeuTrpAspGlnGlnSerAlaGlnLeuLeuPheHisTyrIleGluThrHisArgPheIle	395	
DB	752 GCTCATCGATGAAGACACACAGACTCTCTCCGCCACCTGATGAAGCGTATCATTTATA	811	
QY	400 GluAlaIleArgAlaGlnGlyThrHisValLeuValHisCysGlyMetGlyValSerArg	419	
DB	812 AACAAAGGAAGACGAAACCATTCACAGTGCCTGTGGCTATTCAAAAGGGGGGTAGTTCG	871	
QY	420 SerAlaAlaThrValLeuAlaTyrAlaMetGlyGlnTyrGlyCysSerLeuGlnGlnAla	439	
DB	872 TCGGCTCCACAGTCAATAGCTTATGCATGAAAGGATTTGGCTGGCTCTGGAAGAAAGCA	931	
QY	440 LeuArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGln	455	
DB	932 TATACATCTATGTAAGACAGACAGCGCATTCACGGGCCCAAGCGGGGCTTATAGGCAG	991	
QY	460 LeuGlnIleTyrGlnGlyIleLeuThrAla 469		
DB	992 CTGTCTGATGTAAGAGCATCTTGGATCA 1021		

THIS PAGE BLANK (uspto)

GenCoife version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 04:02:23 ; Search time 159 Seconds
(without alignments)
908.458 Million cell updates/sec

Title: US-09-761-640-4
Perfect score: 2436
Sequence: 1 MALVTSRSPSGASTPVG.....PNDPLRLQIYQILTART 471

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n_model -DEV=xlh
-Q=/cgn2_1/USFPO.spool/US09761640/runtac.08042003_141437_20394/app_query.fasta-1.647
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIG=200 -NORM=ext -THR_SCORE=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09761640.ecgn_1.1.66 -tunat.08042003_141437_20394 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAL -LARGEOUTRY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued Patents_NA.*
- 2: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219.5	9.0	1987	2	US-08-990-379-1
2	216	8.9	531	4	US-09-704-139-3
3	216	8.9	1380	4	US-09-704-139-1
4	216	8.9	1993	2	US-08-990-379-2
5	211	8.7	1691	3	US-09-013-881-12
6	207	8.5	1238	2	US-08-530-290-11
7	194	8.0	912	4	US-09-163-833-1
8	193	7.9	594	4	US-09-163-833-3
9	187	7.7	789	4	US-09-164-193-6
10	187	7.7	789	4	US-09-221-448A-6
11	187	7.7	1016	4	US-09-164-193-4
12	187	7.7	1016	4	US-09-221-448A-4

13	186.5	7.7	861	1	US-07-988-273-1	Sequence 1, Appl1
14	186.5	7.7	861	5	PCT-US93-12019-1	Sequence 1, Appl1
15	185.5	7.6	1358	3	US-09-013-881-11	Sequence 11, Appl1
16	182	7.5	1729	4	US-09-045-973-6	Sequence 6, Appl1
17	182	5.6	821	2	US-08-530-290-9	Sequence 9, Appl1
18	134	5.5	821	2	US-08-933-750C-70	Sequence 70, Appl1
19	134	5.5	821	3	US-09-234-613-70	Sequence 7, Appl1
20	123.5	5.1	7295	2	US-08-487-826B-15	Sequence 15, Appl1
21	113.5	4.7	1098	3	US-09-040-681A-5	Sequence 5, Appl1
22	113.5	4.7	1098	4	US-09-497-897-5	Sequence 2, Appl1
23	110.5	4.5	71989	4	US-09-443-501A-2	Sequence 2, Appl1
24	109.5	4.5	35828	4	US-09-449-218D-17	Sequence 17, Appl1
25	108.5	4.5	80161	3	US-09-036-987A-1	Sequence 1, Appl1
26	108.5	4.5	80161	4	US-09-370-700-1	Sequence 1, Appl1
27	108	4.4	4031	1	US-08-159-784-1	Sequence 1, Appl1
28	107	4.4	237	2	US-08-530-290-5	Sequence 5, Appl1
29	107	4.4	5494	4	US-08-910-864-10	Sequence 10, Appl1
30	107	4.4	5510	1	US-08-123-161A-7	Sequence 7, Appl1
31	107	4.4	5510	4	US-08-483-278-7	Sequence 7, Appl1
32	106.5	4.4	15611	4	US-09-441-340-1	Sequence 1, Appl1
33	106.5	4.4	68750	3	US-09-335-409-1	Sequence 1, Appl1
34	106.5	4.4	68750	4	US-09-568-102-1	Sequence 1, Appl1
35	106.5	4.4	68750	4	US-09-567-969-1	Sequence 1, Appl1
36	106.5	4.4	68750	4	US-09-568-480-1	Sequence 1, Appl1
37	106.5	4.4	68750	4	US-09-568-486-1	Sequence 1, Appl1
38	106.5	4.4	68750	4	US-09-568-472-1	Sequence 1, Appl1
39	106.5	4.4	68750	4	US-09-567-899-1	Sequence 1, Appl1
40	106	4.4	4456	4	US-09-095-443-1	Sequence 1, Appl1
41	105	4.3	50341	1	US-08-247-901C-1	Sequence 1, Appl1
42	105	4.3	50341	2	US-09-075-904-1	Sequence 1, Appl1
43	105	4.3	52297	4	US-09-426-436-1	Sequence 1, Appl1
44	105	4.3	52297	4	US-08-705-557-1	Sequence 1, Appl1
45	103	4.2	3288	2	US-08-620-694A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-990-379-1
Sequence 1, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stork, Phillip J
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
FILE REFERENCE: 4104-00032205A
CURRENT APPLICATION NUMBER: US/08/990.379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000.263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1987
TYPE: DNA
ORGANISM: Rattus norvegicus
US-08-990-379-1

Alignment Scores:

Pred. No.: 1.33e-12 Length: 1987
Score: 219.50 Matches: 112
Percent Similarity: 36.71% Conserves: 51
Best Local Similarity: 25.23% Mismatches: 145
Query Match: 9.01% Indels: 137
DB: 2 Gaps: 17

US-09-761-640-4 (1-471) x US-08-990-379-1 (1-1987)

QY 50 AspGlyAspAsnAspAlaAlaGluAlaSerSerGluProThrGluLysAlaPro 69
||||||| : : : : : |||

Db	386	GATGGAGGAAC	TGGGGAGATGAGACTGCAGCTCTCAAAAGCTATGAA	-----CCG	433
QY	70	SerGIuGIuGLuLeuHISGLYAspGIuThrAspPheGLYcInGLYserGIuSerProGIu			89
Db	440	AGATGAGAAACGGCGGACAGCC	-----GGCAGACAGCGCGGAGCCACGG		484
QY	90	Iys-----GIuGIuGLuGIuArgGIuHISLeuHISLeuMetValGIuLeuLeuArgPro			107
Db	485	CCCCCTGGGGCTGCTGAGCGGGCGCA	-----		511
QY	108	GIuAspAspIleArgLeuAlaAlaGIuLeuGIuAlaProArg			121
Db	512	-----GTGCTTGCTGCTGAGACTGAGGCCCTTGCTGCTACAGCGCGGCTTA			559
QY	122	ProArgValu-----ArgTYrLeuLeuValValSerThrArgGLuGLYcInGLYLeu			139
Db	560	CATCCGAGGCTCGGTGAACATGCGCTGCAATACCATCTGCGCGGAGGAGCCAAAGGCTC			619
QY	140	SerGIuAspGIuThrValLeuLeuGLYValAspPheProAspSerSerProSerCys			159
Db	620	CGTGAG	-----CCTGGACACATTTCTGCCCCGA		649
QY	160	ThrLeuGLYLeuValu-----ProLeuTrpSerAspThrGIuValTYrLeuAspGLY			177
Db	650	GGAAGAGGTGGCGCCCTGCGCTGAGCTGAGCTTACTGCGCTGCTCATCTCATGATGACGA			709
QY	178	AspGLYGIuPheSerValThrSerGIYGIuSerArgIleArgLeuPheIysProIleSerIle			197
Db	710	GCCCGCGCGGAGAGTCTCGG--GGAGACACACAGTCTC			749
QY	198	GIuThrMetTrpAlaThrLeuGIuInValLeuHISGIuAlaCysGIuAlaLeuGLYser			217
Db	750	-----GCTGGTCTGCACAGCTTGGCGCGG			773
QY	218	GIY-----LeuValProGLYIYserAlaLeuThrTrpAlaSer			230
Db	774	AACGGGAGCCACACAGACATCTGCTCTTAAGATGGC	-----		812
QY	231	HisTYrGIuGIuArgLeuAsnSerGIu-----GIuSerCysLeuAsnGIuTrpThrAla			248
Db	813	-----TATGAGAGGTTTCTTCTGATGATCCAGAAATTCTGC--TCTAAACTAAGGCC			863
QY	249	MetaIAspArgLeuIserLeuArgProProSerAlaGIuProGLYIYserSerGIuGIu			268
Db	864	CTGGCGCCATCCACACCCCGTATCCTCCAGACA	-----		899
QY	269	GIuGIuMetGIuAlaIleArgAlaGIuLeuTrpIYsValLeuAspValSerAspLeu			288
Db	899	-----			899
QY	289	GIuSerValThrSerIYsGIuIleArgGIuAlaLeuGIuLeu-----ArgLeuGLY			305
Db	900	-----AATGATCTCTTGATCTGCGCTGCGACGCTCGTGCGG			935
QY	306	LeuProLeuGIuGIuTYrArgAspPheIleAspAsnGIuMetLeuLeuValAlaGIu			325
Db	936	ACCCCACTGCACGACGAGGGGGGCTCGTGGAG	-----		968
QY	326	ArgAspArgAlaSerArgIlePheProHISLeuTYrLeuGLYserGIuTrpAsnAlaIa			345
Db	969	-----ATCTCTCTCTTCTCTTACTGCGGAGAGTCCATCCACACCTGCTCC			1010
QY	346	AsnLeuGIuGIuLeuGIuArgAsnArgValThrHISIleAsnMetAlaArgIuIle			365
Db	1011	CGCAGGAGCACTTATAGCCCTGGGAGATCAGCGCTCTACTGAATGTCTCTCAAGATGCG			1070
QY	366	AspAsnPheTYrProGIuArgPheThrTYrHISAsnValArgLeuTrpAspGIuIuSer			385
Db	1071	CCCATCATCTTGTAGGAGCAATTACCAAGTACCAATGCATCCGGTGAAGATTAACCAAG			1130
QY	386	AlaGIuLeuLeuProHISTrpIYsGIuThrHISArgPheIleGIuAlaIaArgAlaGIu			405
Db	1131	GCTAGACTACACTCTCGTTCTATGAGAGCCATCGAATGATATAGCAGATGAAGACTGC			1190

QY	406	GLYTHHisValIeuValHisCysLysMetGIYValSerArgSerAlaAlaThrValIeu	425
Db	1191	CGAGGGCGAGCTGGTTCACGTCCAGCGCCGCACCTCTAGATCCAGCCACCATCTGCCTG	1250
QY	426	AlATyAlAMetLysGIYThrLysCysSerLeuGIuAlAlaLeuArgHisValGIu	445
Db	1251	GGCTACCTATATATGAAGAAGCGGGTGGAGAGCTTGGAGAGAGCTTTGGAGTGTCTCAAGACAG	1310
QY	446	LeuArgProIleAlaArgProAsnProGIYPhleuArgGIu---LeuGIuIleTyGIu	464
Db	1311	CGCGGTGAGCATCATCTCGCCCAACTTCAGCTTCATGAGGCCAGTGTCTGCAGTTCAGTCT	1370
QY	465	GLYIleLeuThr 468	
Db	1371	CAGGTGCTCACC 1382	
RESULT 2			
US-09-704-139-3			
; Sequence 3, Application US/09704139			
; Patent No. 6420153			
; GENERAL INFORMATION:			
; APPLICANT: Kapeller, Rosanna			
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR			
; FILE REFERENCE: 10448-018001			
; CURRENT APPLICATION NUMBER: US/09/704.139			
; PRIORITY FILING DATE: 2000-11-01			
; PRIOR APPLICATION NUMBER: US 60/185,772			
; PRIORITY FILING DATE: 2000-02-29			
; NUMBER OF SEQ ID NOS: 5			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 3			
; LENGTH: 531			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-704-139-3			
Alignment Scores:			
Pred. No.:		3,77e-13	Length: 531
Score:		216.00	Matches: 55
Percent Similarity:		54.61%	Conservative: 22
Best Local Similarity:		39.01%	Mismatches: 64
Query Match:		8.87%	Indels: 0
DB:		4	Gaps: 0
US-09-761-640-4 (1-471) x US-09-704-139-3 (1-531)			
QY	331	ArgIlePheProHisLeuTyLeuGIYSerGIuThrAsnAlaAlaAsnLeuGIuIeu	350
Db	61	CGCGTCGGCGCCCTACCTCTCTCGGGAGCGCGGAGCGCGGGCGGAGGACAGCTG	120
QY	351	GIuArgAsnArgValThrHisIleLeuAsnMetAlaArgGIuIleAspAsnPhetyrPro	370
Db	121	GGCGCGCGGGAGTACACCTGTGCCTCAACGTTCGCCCCAGCAGCGCGCGCGCGCG	180
QY	371	GIuArgPheThrTyHisAsnValArgLeuThrAspGIuIleSerAlaGIuIleLeuPro	390
Db	181	CGCGCGGTGGCAGAGCTGCCGTGCCGTGTTCACAGACCGCGGTGAGAGACTCTGGCG	240
QY	391	HisTrpLysGIuThrHisArgPheIleGIuAlaAlaArgAlaGIuIleThrHisValIeu	410
Db	241	CACGTGAGCCACACCTGGCGCGCATGGAGCGCGGGTGGCGCGCGCGCGCTGCTA	300
QY	411	ValHisCysLysMetGIuValSerArgSerAlaAlaThrValIleuAlaTyAlaMetLys	430
Db	301	GTCTACTGTCAAGAACGGCGGAGCGCGCTCGCGCCGCTGTCCACCGGTACTCTGCGG	360
QY	431	GIuTyrgLysSerLeuGIuAlaLeuArgHisValGIuIleuArgProIleAla	450
Db	361	CACCGCGGCTCAGCCTGGCAGAGCCTTCACAGATGTGAAGAGCCTCGCCCGGTAGCA	420
QY	451	ArgProAsnProGIYPhleuArgGIuIleuGIuIleTyArgGIuIleuThrAlaArg	470


```

Db 581 GCGCTG-----586
Qy 245 GIUTRPTThralametalaspleuuserleuargproproseralagluiproglygly 264
Db 587 CATATACATCGTGGGGGGGGGCAAGGGCTCGTAGCTTGACACAGAT-----637
Qy 265 serseglu-glnlglnmetgluInalalAargalagluLeuTrpysValleuas 284
Db 638 TCTCCCCCGGAGAGAGAGTCCCGCCCGCTCGGCTCTACTCGGCTGTCAAT 637
Qy 284 pValserAspLeuGluSerValThrSerlysgluLeuargGlnalaleuGluLeuArgle 304
Db 698 CGCTCGATGAGCCGAGCCCGCCGAGAGCTCCGGGAG-----GACAGCACAGT 751
Qy 304 uGlyLeuProLeuGlnGlnIlyTargAspPheLeasp---AsnGlnMetLeuLeuVa 323
Db 752 GTCGCTGTCGTGACAGCGCTGGCGGAGACGGAGCCAGACATCTGCTGCTTAA 811
Qy 323 lAlaglnArgAspArgAlaser-----330
Db 812 AGGTGCTATGAGAGGTTTCTCTGTAGTACCCAGATTTGCTTAAACTAAGGCCCT 871
Qy 330 -----330
Db 872 GCGCGCCATCCACCCCGCTACCTCCAGACAAATGAGTCTGTGATCGGGCTGCAG 931
Qy 331 -----ArgllePheProHisLeuTy 337
Db 932 CTCCTGTGGAGCCCACTGCAGACACAGGGGGTCTGTGAGATCTCTTCTCTTA 991
Qy 337 rLeuGlySerGluTrpAsnAlaAlaAsnLeuGluGluGlnArgAsnArgValThrnl 357
Db 992 CCGCGAGTGCCTTACACGCTGCGCGAGGAGCATGCTTGATGCCCTGGGAGTACAGGC 1051
Qy 357 sIleLeuAsnMetAlaArgGluLeuAspAsnPhetYrProGluArgPheThrYrHisAs 377
Db 1052 TCATCTGATGTCTCTCAACATGCCCCCACTTGAGGAGCATTTACAGTCAAGTG 1111
Qy 377 nValArgLeuTrpAspGluLeuSerAlaGlnLeuLeuProHisTrpLysGluThrHisAr 397
Db 1112 CACCGCGTAGAGATACCAACAGGCTGACATCAGCTCTGTTTCATGAGAACCATCGA 1171
Qy 397 gPheIleGluAlaAlaArgAlaGlnIlyThrHisValLeuValHisCysLysMetGlyVa 417
Db 1172 ATCATATGACGAGTGAAGAGACTCCGAGGGGAGTGTGCTGCTGCTGCTGCTGCTG 1231
Qy 417 lSerArgSerAlaAlaThrValLeuAlaTyralaMetLysGlnTyrgLysSerLeuGl 437
Db 1232 CTCTAGATCAGCCACATCTGCTGGCTTACCTGATGATGAAGAAACGGGTGAGGCTGGA 1291
Qy 437 uGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLe 457
Db 1292 GAGAGGCTTTCGATGTTGTCAGACAGCGCGCTGACATCATTCGCCCAACTTCAAGCTTCA 1351
Qy 457 uArgGln---LeuGlnIleTyrglnGlyLeuLeuThr 468
Db 1352 GGGCAGAGTGTGCTGAGTTCGATCTCAGGTCTCAGC 1388

```

RESULT 5
US-09-013-881-12
Sequence 12, Application US/09013881

GENERAL INFORMATION:
PATENT NO. 6132964
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.

```

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,881
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0470 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1691 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: MUSCNOT02
CLONE: 971204
US-09-013-881-12
Alignment Scores:
Pred. No.: 7,56e-12 Length: 1691
Score: 211.00 Matches: 56
Percent Similarity: 56.46% Conservative: 27
Best Local Similarity: 38.10% Mismatches: 56
Query Match: 8.66% Indels: 8
DB: 3 Gaps: 4
US-09-761-640-4 (1-471) x US-09-013-881-12 (1-1691)
Qy 327 AspArgAlaserArgllePheProHisLeuTyrgLysGluTrpAsnAlaAlaAsn 346
Db 589 AACCATGCGCAGAGGCTGGCCAGGCTCTATCTGGAGACAGACATGGCTAACAA 648
Qy 347 LeuGluGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAsp 366
Db 649 CGCGGAGAGCTTGGCGCGCTGGGCAATCAGACAGCTCTCATATGCTGAGAGAGAG 708
Qy 367 AsnPhetYrProGluArgPheThrYrHisAsnValArg---LeuTrpAspGluLeuSer 385
Db 709 CGAGGACAGCGCCAGAGGCTGATGAGGGGCTGGGATCGGCTGAGGCTGTGAGCCAG 768
Qy 386 AlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIleGlnAlaAlaArgAlaGln 405
Db 769 TTTGACATGAGATCCACTTCCAGAGCGGCTGCGGACTTATCCACCGGCGGCTGAGCC 828
Qy 406 ---GlyThrHisValLeuValHisCysLysMetGlyValAsnArgSerAlaAlaThrVal 424
Db 829 CAGAGAGGAGAGATCCGTGGTGGTGTGGTGGGCGGTGAGAGAGAGAGAGAGAGAGAG 888
Qy 425 LeuAlaTyralaMetLysGlnTyrgLysSerLeuGluGlnAlaLeuArgHisValGln 444
Db 889 CTGGCTACCTCATGCTGTGACACACACCTTACCTGCTGAGGAGCCATCAAGAAAGTCA 948
Qy 445 GluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu-----460
Db 949 GAGCACCAGAGCATATC---CCCAACCGGGGCTTCTGAGGAGCTCTGGCCCTTGAG 1005

```


ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(789)
US-09-164-193-4

Alignment Scores:

Pred. No.:	9,37e-10	Length:	1016
Score:	187.00	Matches:	47
Percent Similarity:	53.03%	Conservative:	23
Best Local Similarity:	35.61%	Mismatches:	56
Query Match:	7.68%	Indels:	6
DB:	4	Gaps:	2

US-09-761-640-4 (1-471) x US-09-164-193-4 (1-1016)

OY 327 ASPARGALSERARGILEPHEPROHISLEUTYRLEUGLYSERGLUTRPSNALAALAASN 346
DB 49 AACCATGCCGACGAGGTCTGGCCAGGCGCTTCTATCTCGAGACCAAGACATGGCTAACAC 108
OY 347 LEUGLULLEUGLINAARGVVALTHRHISILEUASMETALAARGLUILEASP 366
DB 109 CGCCGGAGCTTCGCCGCTGGGCATACGACAGTCTCAATGCCACACAGCCGGTGG 168
OY 367 ASNPHEPTYRPROGLUARGPHE-----THYTHISASNYVALARGLEUTRP 381
DB 169 CGAGCAGCCCGACGAGGCTATGAGGGCTGGGCATCCGCTACCTGGGTGTGAGCCGAC 228
OY 382 ASPGLUGLUSERALAGLNEUPROHISTRPLYSGLUTHISARGPHEILEGUALA 401
DB 229 GACTGCCAGCGCTTGACATGAGCATCTCCAGACGGCTGCCAGCTTCATCCACCGG 288
OY 402 ALAARGALAGL--GLYTHRHISVALLEUVALHISCYLSYMETGLYALSERARGSER 420
DB 289 GCGCTAGCCGACGAGGAGGAAGATCCTGTCATCTGCTGTGGCGGTGAGCCGATCC 348
OY 421 ALAALATHVALLEUALATYRALAMETLYSGINTYRGLUCYSERLEUGLINALALEU 440
DB 349 GCCACCTGTACTGGCTACCTCATCTGTACACACACCTTACCTCGTGGAGGCCATC 408
OY 441 ARGHISVALGLNGLLEUARGPROILEALARGPRO 452
DB 409 AAGAAAGTCAAGACCAACGAGGAGGAGGCCGAGCC 444

RESULT 12

US-09-221-448A-4
Sequence 4, Application US/09221448A
Patent No. 6436685
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPTP PROTEIN MOLECULES AND USES THEREFOR (AS AMENDED)
FILE REFERENCE: MN1-051DV1
CURRENT APPLICATION NUMBER: US/09/221,448A
CURRENT FILING DATE: 1998-12-28
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 4
LENGTH: 1016
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(789)
US-09-221-448A-4

Alignment Scores:

Pred. No.:	9,37e-10	Length:	1016
Score:	187.00	Matches:	47
Percent Similarity:	53.03%	Conservative:	23
Best Local Similarity:	35.61%	Mismatches:	56
Query Match:	7.68%	Indels:	6
DB:	4	Gaps:	2

US-09-761-640-4 (1-471) x US-09-221-448A-4 (1-1016)

OY 327 ASPARGALSERARGILEPHEPROHISLEUTYRLEUGLYSERGLUTRPSNALAALAASN 346
DB 49 AACCATGCCGACGAGGTCTGGCCAGGCGCTTCTATCTCGAGACCAAGACATGGCTAACAC 108
OY 347 LEUGLULLEUGLINAARGVVALTHRHISILEUASMETALAARGLUILEASP 366
DB 109 CGCCGGAGCTTCGCCGCTGGGCATACGACAGTCTCAATGCCACACAGCCGGTGG 168
OY 367 ASNPHEPTYRPROGLUARGPHE-----THYTHISASNYVALARGLEUTRP 381
DB 169 CGAGCAGCCCGACGAGGCTATGAGGGCTGGGCATCCGCTACCTGGGTGTGAGCCGAC 228
OY 382 ASPGLUGLUSERALAGLNEUPROHISTRPLYSGLUTHISARGPHEILEGUALA 401
DB 229 GACTGCCAGCGCTTGACATGAGCATCTCCAGACGGCTGCCAGCTTCATCCACCGG 288
OY 402 ALAARGALAGL--GLYTHRHISVALLEUVALHISCYLSYMETGLYALSERARGSER 420
DB 289 GCGCTAGCCGACGAGGAGGAAGATCCTGTCATCTGCTGTGGCGGTGAGCCGATCC 348
OY 421 ALAALATHVALLEUALATYRALAMETLYSGINTYRGLUCYSERLEUGLINALALEU 440
DB 349 GCCACCTGTACTGGCTACCTCATCTGTACACACACCTTACCTCGTGGAGGCCATC 408
OY 441 ARGHISVALGLNGLLEUARGPROILEALARGPRO 452
DB 409 AAGAAAGTCAAGACCAACGAGGAGGAGGCCGAGCC 444

RESULT 13

US-07-988-273-1
Sequence 1, Application US/07988273
Patent No. 5512434
GENERAL INFORMATION:
APPLICANT: ARMONSON, Stuart A.
APPLICANT: BOTTARO, Donald P.
APPLICANT: ISHIBASHI, Toshio
APPLICANT: MIKI, Toru
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988, 273
FILING DATE: 19921214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/182 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS

```

US-07-988-273-1      LOCATION: 29...586 ?
Alignment Scores:
Pred. No.:           8,11e-10          Length: 861
Score:              186.50             Matches: 51
Percent Similarity: 49.09%             Conserved: 30
Best Local Similarity: 30.91%          Mismatches: 73
Query Match:        7.66%               Indels: 11
DB:                  1                   Gaps: 4

US-09-761-640-4 (1-471) x US-07-988-273-1 (1-861)
OY   306 LeuProLeuGInGInIYrArGaSPheIIeaSprsnImIetLeuLeuValAlaGln 325
    ||| :||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   47 CTCtCGGGCAGAGATCTCAACAGACTGTCTCGACGGCAGACGGCTCTACAGCTCCG 106
OY   326 ArgAspArgAlaSerArgGIlePheProHisIleuTYrLeuGIySerGIutPraAlaIa 345
    :: :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   107 AGCCAGCCCTCAACAGAGAGTACCSCCGSGATCTACGTGGGCAACGGCTGTGGCTCAG 166
OY   346 AsnLeuGInGInLeuGInAraGAsnArgValThrhHisIleLeuSnmEtAlaArgIuIle 365
    :::::||||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db   167 GACATCCCCAAGCTGCAGAAACTAGCATGCCATCCCATGTGCTGAACCGCGTGAGGCGAG 226
OY   366 Asp-----AsnPheTYrProGluArg--PheThrTYrHis 376
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   227 TCCTTCATGCACGTCAACACCAATGCCAATCTCTCACAAAGACTCGGCATCACATACCTG 286
OY   377 AsnValArgLeuTrpAspGIuSerGIuSerAlaGInLeuLeuProHisIstrPIlySIuPrHris 396
    ::::: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   287 GGCAATCAAGGCCAACGACACACAGAGATTCAACCTCAGCGCTTAATTGAAAGGGCTGCC 346
OY   397 ArGPHeIIeGInAlaAlaArgAlaGln---GIYThrhISValLeuValHIscysLysMet 415
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   347 GACTTCATATGACAGAGCTTTGGCTCAAAAGAATGCGCGGGTCTCTGCTCCACTGCCGGAA 406
OY   416 GIYalSerArgSerAlaAlaThrhValLeuAlaTYralaMetLysGInTYrGIuCySer 435
    ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db   407 GGTATAGCCCTCCCTCCCAACGCTGTATATGCTGACCTCATGTAGTGGCGAGAGATGGAC 466
OY   436 LeuGInGInAlaLeuArGHisValGInGInLeuArprroIleaIarqProdanProGIy 455
    ::::: ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db   467 GTCAAGTCTGCCCTGAGACATCGTAGGACAGAACCGTGAATCGGC---CCAAAGATGGC 523
OY   456 PheLeuArgInLeu 460
    ||||| |||||
Db   524 TTCCTGGCCAGCTC 538

RESULT 14
PCT-US93-12019-1
Sequence 1, Application PC/TUS9312019
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MYS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12019
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,273
FILING DATE: 14-Dec-1992
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 861 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

?	TOPOLOGY:	Linear		
?	FEATURE:			
?	NAME/KEY:	CDS		
?	LOCATION:	29..586		
?	PCT-US93-12019-1			
	Alignment Scores:			
	Pred. No.:	8,11e-10	Length:	861
	Score:	186.50	Matches:	51
	Percent Similarity:	49.09%	Conservative:	30
	Best Local Similarity:	30.91%	Mismatches:	73
	Query Match:	7.66%	Indels:	11
	DB:	5	Gaps:	4
	US-09-761-640-4 (1-471) x PCT-US93-12019-1 (1-861)			
QY	306	LeuPLeuEngInGlnTYrArGAspPheIleAspAsnGlnMetLeuLeuValAlaGln	325	
		:::	:::	
Db	47	CTCTCGGCGGAGATCTCAACAGACCTCTCTCGGAGCGGAGCGCTCTCAGACCTCCG	106	
QY	326	ArgAspArgAlaSerArgIlePheProHisLeuTYrLeuGlySerGluTYrPAsnAlaAla	345	
		:::	::: : :	
Db	107	AGCCAGCCCTCAGACGAGGTACCCGCGGATCTAGCTGGGCAACGGCTGTGGCTCAG	166	
QY	346	AsnLeuGlnGluGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIle	365	
		:::	::: :	
Db	167	GACATCCCCAGCTCGAGAAACTAGGATCATCCCATGTGCTGAAAGGGGCGTAGGCGCAGG	228	
QY	366	Asp-----AsnPheTYrProGluArg--PheThrTYrHis	376	
		:::	:::	
Db	227	TTCCTTCATGCACGTCAACACCAATGCCAATGCCAATCTTCAAGAGACTCCGACATACATACCTG	286	
QY	377	AsnValArgLeuTYrPAspGluGluSerAlaGlnLeuLeuProHisIleTYrPAsnThrHis	396	
		:::	::: :	
Db	287	GGCATCAAGGCCACAGCAGACACAGAGATTCACCTCAGCGCTTATGAAAGGCGTCGC	346	
QY	397	ArgPheIleGluAlaAlaArgAlaGln--GlyThrHisValLeuValHisCysLysMet	415	
		:::	::: :	
Db	347	GACTTCATGTGACGAGCGTTTGCTCAAAAGATGCGCGGTGCTGCTCCACTCCGGGA	406	
QY	416	GlyValSerArgSerAlaAlaThrValLeuAlaTYrAlaMetLysGlnTYrGluCysSer	435	
		:::	::: :	
Db	407	GGTATATACCCCTCCGCCACAGCGCTAGTATCGCTACCTCATGATGCGGACAGAGATGCAC	466	
QY	436	LeuGlnGluAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGly	455	
		:::	::: :	
Db	467	GTCAGAGTGGCTGCCAGAGATGCTAGGAGCAAGACCGTGAGATCGCG--CCCAAGATGGC	523	
QY	456	PheLeuArgGlnLeu	460	
		:::	:::	
Db	524	TTCCTGGCCGACGCTC	538	
	RESULT 15			
	US-09-013-881-11			
	Sequence 11, Application US/09013881			
	Patent No. 6132864			
	GENERAL INFORMATION:			
	APPLICANT: Bandman, Olga			
	APPLICANT: Lal, Preeti			
	APPLICANT: Hillman, Jennifer L.			
	APPLICANT: Corley, Neil C.			
	APPLICANT: Gugler, Karl J.			
	APPLICANT: Shah, Purvi			
	TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES			
	NUMBER OF SEQUENCES: 16			
	CORRESPONDENCE ADDRESS:			
	ADDRESSEE: Incyte Pharmaceuticals, Inc.			
	STREET: 3174 Porter Drive			
	CITY: Palo Alto			
	STATE: CA			
	COUNTRY: USA			
	ZIP: 94304			
	COMPUTER READABLE FORM:			

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013.881
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0470 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TMLR3DT02
CLONE: 507537
US-09-013-881-11

Alignment Scores:
Pred. No.: 2.1e-09 Length: 1358
Score: 185.50 Matches: 63
Percent Similarity: 43.18% Conservative: 32
Best Local Similarity: 28.64% Mismatches: 98
Query Match: 7.61% Indels: 27
DB: 3 Gaps: 4

US-09-761-640-4 (1-471) x US-09-013-881-11 (1-1358)

QY 256 ArgProProSerAlaGluProGlyGlySerSerGluGlnGlnMetGluAlaIle 275
DB 9 CGTCCGAGAGCCGCGAGCTGTCTGTGGCGGCCCATGTGGAGGAGGAT-----GTT 53
QY 276 ArgAlaGluLeuTrpLysValLeuAspValSerAspLeuGluSerValThrSerLysGlu 295
DB 54 CTAGCCGGAGTCTACGATG----- 74
QY 296 IleArgGlnAlaLeuGluLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPhe 315
DB 75 GTAGGCGCAGAAAGCCGCTGTCTGTGGCGGCCCATGTGGAGGCTCCGGGCCGAGT 134
QY 316 Asp-----AsnGlnMetLeuLeuValAlaGlnArgAspArgAlaSer 330
DB 135 GATGGCTGCGACCTAGCAACCCAGCCAGCAGACAGTCAGCTGCCGGCGATGCTG 194
QY 331 ArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaAlaAsnLeuGluLeu 350
DB 195 GAAGTGCACACGAGATTGATTGCGTGGGCGCGCGCGCGGAGGAGGATCAGTCTG 254
QY 351 GlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrPro 370
DB 255 ACGGAAGCGCGGATCAGCGCCGCTGTACAGTGCAGTCCGAGAGCCAGCTCAAGCGC 314
QY 371 GluArgPheThrTyrHisAsnValArgLeuTrp-----AspGlnGluSerAla 386
DB 315 GGGCGTGGGGTGAAGATCTATGGCGCTCTTGTGTCACAGCGCTGACCAACCCGAGACG 374
QY 387 GlnLeuLeuProHisIleTrpLysGluThrHisArgPheIleGluAlaIleArgAlaGlnGly 406
DB 375 GAACCTACTCAGCATCTGAGCCGCGTCCGCGCTTTCATCGGTACGAGCCGCGGTAGAGGC 434
QY 407 ThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaIleThrValLeuAla 426

DB 435 CGAGCGTGTGTGGTGCACCTGTATGCGAGAGATCGAAGTGTGGCCATATTAATCTGCT 494
QY 427 TyrAlaMetLysGlnTyrGluCysSerLeuGlnGlnAlaLeuArgHisValGlnGluLeu 446
DB 495 TTTCATGTAAGACTGACCACTTCCCTTTGAAGAAAGCCTATGAAAGGCTCCAGATTCTC 554
QY 447 ArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnGlyLe 466
DB 555 AAACGAGCGCTTACATGATGAGGGGTTTGAGTGGCAACTGAATTTATCCAGGCAATG 614

Search completed: April 12, 2003, 10:55:32
Job time : 176 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 11, 2003, 03:44:27 ; Search time 3860 Seconds

(without alignments)
1976.185 Million cell updates/sec

Title: US-09-761-640-4
Perfect score: 2436
Sequence: 1 MALVTVSRRSPGSGASTPVG.....PNEGFLRLQIYQILTRAT 471

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=x1h
-O=/cgn2_1/USPFO.spool/US0976140/rnact_08042003_141435_20370/app_query.fasta.1.647
-DB=EST -QPM=fastcap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09761640_@CGC_1_1_3256_@runatc_08042003_141435_20370 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAB -LARGEDUERY -NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THREHDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%		Query		Match		Length		DB		ID		Description	
No.	Score	Match	Length	DB	ID	No.	Score	Match	Length	DB	ID	No.	Score	Match	Length
1	1413	58.0	1004	14	BM910493	1	1413	58.0	1004	14	BM910493	1	1413	58.0	1004
2	1350	55.4	991	13	BM471207	2	1350	55.4	991	13	BM471207	2	1350	55.4	991
3	1325	54.4	875	12	BG752969	3	1325	54.4	875	12	BG752969	3	1325	54.4	875
4	1318	54.1	1024	13	BM560192	4	1318	54.1	1024	13	BM560192	4	1318	54.1	1024
5	1301.5	53.4	812	12	BG750408	5	1301.5	53.4	812	12	BG750408	5	1301.5	53.4	812
6	1202	49.3	790	12	BG326709	6	1202	49.3	790	12	BG326709	6	1202	49.3	790
7	1147.5	47.1	846	10	BE563259	7	1147.5	47.1	846	10	BE563259	7	1147.5	47.1	846
8	1147.5	47.0	910	12	BG169297	8	1147.5	47.0	910	12	BG169297	8	1147.5	47.0	910
9	1144.5	47.0	726	12	BE869222	9	1144.5	47.0	726	12	BE869222	9	1144.5	47.0	726
10	1091.5	44.8	919	12	BE737087	10	1091.5	44.8	919	12	BE737087	10	1091.5	44.8	919
11	1089.5	44.7	943	13	BM563401	11	1089.5	44.7	943	13	BM563401	11	1089.5	44.7	943
12	1085	44.5	932	12	BG340220	12	1085	44.5	932	12	BG340220	12	1085	44.5	932
13	1042.5	42.8	897	12	BE873337	13	1042.5	42.8	897	12	BE873337	13	1042.5	42.8	897
14	1009	41.4	663	12	BG753072	14	1009	41.4	663	12	BG753072	14	1009	41.4	663
15	989.5	40.6	716	12	BE907514	15	989.5	40.6	716	12	BE907514	15	989.5	40.6	716
16	982	40.3	803	13	BI105472	16	982	40.3	803	13	BI105472	16	982	40.3	803
17	968	39.7	589	9	AL602806	17	968	39.7	589	9	AL602806	17	968	39.7	589
18	966	39.7	791	12	BE881145	18	966	39.7	791	12	BE881145	18	966	39.7	791
19	959.5	39.4	892	12	BF971565	19	959.5	39.4	892	12	BF971565	19	959.5	39.4	892
20	916.5	37.6	1228	14	BM921421	20	916.5	37.6	1228	14	BM921421	20	916.5	37.6	1228
21	910	37.4	969	10	BE563116	21	910	37.4	969	10	BE563116	21	910	37.4	969
22	900.5	37.0	704	14	BQ108910	22	900.5	37.0	704	14	BQ108910	22	900.5	37.0	704
23	899	36.9	846	13	BI820326	23	899	36.9	846	13	BI820326	23	899	36.9	846
24	897	36.8	937	14	BQ718968	24	897	36.8	937	14	BQ718968	24	897	36.8	937
25	884.5	36.3	658	12	BG822776	25	884.5	36.3	658	12	BG822776	25	884.5	36.3	658
26	870	35.7	863	12	BG749256	26	870	35.7	863	12	BG749256	26	870	35.7	863
27	805	33.0	1120	12	BI155292	27	805	33.0	1120	12	BI155292	27	805	33.0	1120
28	797.5	32.7	860	12	BE127735	28	797.5	32.7	860	12	BE127735	28	797.5	32.7	860
29	770	31.6	652	10	BB660625	29	770	31.6	652	10	BB660625	29	770	31.6	652
30	767	31.5	612	13	BM017088	30	767	31.5	612	13	BM017088	30	767	31.5	612
31	767	31.5	1121	13	BM470627	31	767	31.5	1121	13	BM470627	31	767	31.5	1121
32	755.5	31.0	761	12	BF579815	32	755.5	31.0	761	12	BF579815	32	755.5	31.0	761
33	728	29.9	471	13	BM482726	33	728	29.9	471	13	BM482726	33	728	29.9	471
34	702	28.8	424	10	AW326161	34	702	28.8	424	10	AW326161	34	702	28.8	424
35	691	28.4	1125	14	BM809903	35	691	28.4	1125	14	BM809903	35	691	28.4	1125
36	689	28.3	599	10	BE280203	36	689	28.3	599	10	BE280203	36	689	28.3	599
37	689	28.3	1069	13	BM560223	37	689	28.3	1069	13	BM560223	37	689	28.3	1069
38	688	28.2	743	13	BM017302	38	688	28.2	743	13	BM017302	38	688	28.2	743
39	687	28.2	409	10	AW732634	39	687	28.2	409	10	AW732634	39	687	28.2	409
40	684.5	28.1	771	13	BI146038	40	684.5	28.1	771	13	BI146038	40	684.5	28.1	771
41	684	28.1	1158	14	BM910866	41	684	28.1	1158	14	BM910866	41	684	28.1	1158
42	672	27.6	799	13	BG965294	42	672	27.6	799	13	BG965294	42	672	27.6	799
43	662	27.2	629	14	BM785529	43	662	27.2	629	14	BM785529	43	662	27.2	629
44	641	26.3	890	12	BE901875	44	641	26.3	890	12	BE901875	44	641	26.3	890
45	631.5	25.9	986	14	BQ715279	45	631.5	25.9	986	14	BQ715279	45	631.5	25.9	986

ALIGNMENTS

RESULT 1
LOCUS BM910493 1004 bp mRNA linear EST 12-MAR-2002
DEFINITION ACENCCOURT 6703004 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5455997
5' mRNA sequence.
ACCESSION BM910493
VERSION BM910493.1 GI:19360872
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1004)
NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM1948 row: 9 column: 22
 High quality sequence stop: 725.
 Location/Qualifiers

FEATURES

source

1. 1004
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5453997"
 /clone_1lb="NIH_MGC_98"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 211 a 299 c 324 g 167 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 2,54e-137 Length: 1004
 Score: 1413.00 Matches: 299
 Percent Similarity: 77.69% Conservative: 4
 Best Local Similarity: 76.67% Mismatches: 18
 Query Match: 58.00% Indels: 70
 Gaps: 4

US-09-761-640-4 (1-471) x BM910493 (1-1004)

QY 25 AlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaValLeuArgGlyAla 44
 DB 2 GGGGTCCAGCAAGAGAGTGCCTCCAGAGGACAGCTTGGCTGCTCCGTGGGCT 61
 QY 45 ValLeuGlyLeuGlnAspGlyAspAsnAspAlaAlaGlnAlaSerSerGluPro 64
 DB 62 GTCTCGGAGCTGAGATGGAGGGGACATGATGATGACAGAGCCAGTTCTGAGCCA 121
 QY 65 ThrGluGlyAlaProSerGluGluGluLeuHisGlyAspGlnThrAspPheGlyGlnGly 84
 DB 122 ACAGAGAAAGGCCCGAGTGTAGAGAGAGCTCCACGGGACACAGACTTCGGGACAGGA 181
 QY 85 SerGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 104
 DB 182 TCCAGAGAGTCCCGAG 241
 QY 105 LeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArg 124
 DB 242 CTGAGAGCCGACAGATGACATCCGCTGGAGAGCCAGCTGAGAGACCCCGGCTCCCGCG 301
 QY 125 LeuArgTyrLeuLeuValValSerThrArgGluGlyGluGlyLeuSerGlnAspGluThr 144
 DB 302 CTCCCTACTCTGCTGATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
 QY 145 ValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuVal 164
 DB 362 GTCTCTCTGGGGGTGATTTCTTCCACAGACACCTCCCGAGCTGACACCTGGGGCTGGTTC 421
 QY 165 LeuProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThr 184
 DB 422 TTGCCCTCTGTGAGTACACCCAGGTGTAATTAGATGGAGAGGGGGCTTCAGCGTACGAG 481

QY 185 SerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMetTrpAlaThrLeu 204
 DB 482 TCTGTGGGCAAGCCGAGTCTTCAAGCCATCTCCATCCAGACATGGG----- 532
 QY 205 GlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySer 224
 DB 532 ----- 532
 QY 225 AlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluGlnSerCysLeuAsn 244
 DB 532 ----- 532
 QY 245 GluTrpThrAlaMetAlaAspLeuGluSerLeuArgProProSerAlaGluProGlyGly 264
 DB 532 ----- 532
 QY 265 SerSerGluGlnGlnGlnMetGlnGlnAlaIleArgAlaGluLeuTrpLysValLeuAsp 284
 DB 533 TCTCAGACAGAGAGAGATGAG 592
 QY 285 ValSerAspLeuGlnSerValThrSerLysGluIleArgGlnAlaLeuGluLeuArgLeu 304
 DB 593 GTCTAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGACAGCTCTGGAGCTGGCCCTG 652
 QY 305 GlyLeuProLeuGlnGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAla 324
 DB 653 GGGCTCCCCCTCAGAGATGACCTGACTTCTATCGACACACAGATGCTGCTGCTGGTGGCA 712
 QY 325 GlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAla 344
 DB 713 CAGCGGACCCGAGCCCTCCGACATCTTCCCTACCTGAGGCTCAGAGTGAACAGCA 772
 QY 345 AlaAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 364
 DB 773 GCAGACCTGGAGAGCTGAG 832
 QY 364 uIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGlu 384
 DB 833 GATTGACACTTCTACCTCGAGGGCTTCACTACACACAG-GTGGGCTTCCGTGGGAGTGA 891
 QY 384 uSerAla-----GlnLeuLeuProHisTrp---LysGluTrpHisArg-PheIleGlnAla 401
 DB 892 AGGAGTGGGGCCAGCTGCTGGCCCGACATGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951
 QY 401 la---AlaArgAlaGlnGlyThrHis 408
 DB 952 GCCTTCCAGAGACACAGAGGCACAC 977
 RESULT 2
 BM471207 991 bp mRNA linear EST 05-FEB-2002
 LOCUS
 DEFINITION AGENCOURT.6478155 NIH_MGC_72 Homo sapiens cdna clone IMAGE:5563111
 5', mRNA sequence.
 ACCESSION
 VERSION BM471207
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS NIH-MGC <http://imgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: L1AM1293 row: j column: 08
 High quality sequence stop: 592.
 Location/Qualifiers
 1. .991
 source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5563111"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 199 a 303 c 329 g 159 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1e-130 Length: 991
 Score: 1350.00 Matches: 278
 Percent Similarity: 91.26% Conservative: 4
 Best Local Similarity: 89.97% Mismatches: 23
 Query Match: 55.42% Indels: 4
 DB: 13 Gaps: 2

US-09-761-640-4 (1-471) x BM471207 (1-991)

QY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
 Db 55 ATGGCCCTGGTGACAGTGAGCCGTTCCGCCCGGCGACGGCGCTCCACGCCGTGGGG 114
 QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgGlnSerPheAlaVal 40
 Db 115 CCCCGGACCGGGGGGTCCACGAGAGTGCATCCAGCCAGGACAGCTTTCGGTGC 174
 QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGluAla 60
 Db 175 CTCGGTGGGCTGTCCTGGGACTCAGAGATGAGAGGACATGATGATGACAGACAGCC 234
 QY 61 SerSerGluProThrGlnGluAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
 Db 235 AGTTCTGAGCCCAAGAGAAAGCCCGGAGTGAAGGAGGCTCCAGGGGACACAGAC 294
 QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
 Db 295 TTGGGGCAGAGATCCCAAGAGTCCCAAGAGAGAGAGACAGACAGCTGCAGCTC 354
 QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGluAlaPro 120
 Db 355 ATGCTACAGCTGTGAGGCCCGCAGAGATGATCCGCTGGCAGCCAGCTGGAGCACC 414
 QY 121 ArgProProArgLeuArgTyrLeuLeuValValSerThrArgGluGluGlyLeuSer 140
 Db 415 CGGGCTCCCGCGGCTGCTACCTCTGCTGTCTTCTACACGAGAAAGAGTCTGAGC 474
 QY 141 GlnAspGlnTrpValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
 Db 475 CAGAGTAGAGAGGCTCTCTCTGGGTGATTTCCCTGACACAGCTCCCGCAGCTGCAC 534
 QY 161 LeuGlyLeuValLeuProLeuTyrSerAspThrGlnValTyrLeuAspGlyAspGly 180
 Db 535 CTGGGCGCTGCTTGGCCCTCTGAGTACACCCAGGAGTGAATGAGAGAGCGGGGC 594
 QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
 Db 595 TTCAGCGGTGAGCTGTGGGCGAAAGCCGAGTCTTCAAGCCCATCTCCAGACCATG 654
 QY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGlyLeu-Val 220
 Db 655 TGGGCGACACTCCAGGTTTTCACCAAGCATGTGAGGAGAGCTCTTAGGAGGCGCTTGTAC 714
 QY 220 LProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGluG 240

Db 715 CCCGGGTGGCAGTGCCCTCAGCTGGGCCAGCCACTACCAAGAGAGACTTGAACTCCGAC 774
 QY 240 nSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAl 260
 Db 775 GAGCTGCTTCATGAGGGAGCGGCTATGGCCGACTGGAGTCTCTGGCGGCTCCGGCCG 834
 QY 260 aGluProGlyGlySerSerGlu-GlnGlnGlnMetGluGlnAlaIleArgAlaGluLeu- 279
 Db 835 CAACCTGGCGGGGTCTCCGAACAGACGAAAGGAGGACGAGCGATCCCTGCTTAACCTG 894
 QY 280 --TrpLysValLeuAspValSerAspLeuGlu--SerValThrSerLysGluIleArg 298
 Db 895 TGGAAAAGGTTTGATTTCTCAGCCCGGAAAGGTCCTCTCCAAAAAACCCTCCG 954
 QY 298 InAlaLeuGluLeuArgLeuGly 305
 Db 955 GGCTTCTGAACTTGCCCTGGGG 977

RESULT 3
 BG752969 875 bp mRNA Linear EST 15-MAY-2001
 LOCUS 602732427F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4876014 5',
 DEFINITION mRNA sequence.
 ACCESSION BG752969
 VERSION BG752969.1 GI:14063622
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 875)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution Information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://Image.lnl.gov
 Plate: L1CM1758 row: e column: 07
 High quality sequence stop: 818.

FEATURES

Location/Qualifiers
 1. .875

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4876014"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. 1"

BASE COUNT 184 a 267 c 275 g 149 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.41e-128 Length: 875
 Score: 1325.00 Matches: 276
 Percent Similarity: 94.86% Conservative: 1
 Best Local Similarity: 94.52% Mismatches: 12
 Query Match: 54.39% Indels: 6
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BG752969 (1-875)


```

|||||
Db 423 TTGCCCTCTGTGAGACACCAGGTACTGATGAGAGAGGGGCTCAGCGTACG 482
QY 185 SerGluGlnSerArgIlePheIleProIleSerIleGlnThrMetTrpAlaThrIleu 204
Db 483 TCTGGTGGGCAAGCGGATGCTTCAAGCCCATCTCCATCCAGACATGTGG----- 533
QY 205 GlnValIleuHisGlnAlaCysGlnAlaAlaLeuGlnSerGlyLeuValProGlyGlySer 224
Db 533 ----- 533
QY 225 AlaLeuThrTrpAlaSerHisTyrGlnGlnIleuAsnSerGlnSerCysLeuAsn 244
Db 533 ----- 533
QY 245 GluThrTrpAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGluProGlyGly 264
Db 533 ----- 533
QY 265 SerSerGluGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpLysValLeuAsp 284
Db 534 TCCCTCAGACAGAGACAGATGAGAGCGCATCCGTGCTGAGCTGTGAAAGTGTGGAT 593
QY 285 ValSerAspLeuGlnSerValThrSerLysGlnIleArgGlnAlaLeuGlnLeuArgLeu 304
Db 594 GTCACTGACCTGAGAGAGTCACTTCCAAAGAGATCCCGCAGGCTCTGGAGCTGGCGCTG 653
QY 305 GlnLeuProIleGlnGlnIleThrArgAspPheIleAspAsnGlnMetLeuLeuValAla 324
Db 654 GGGTCTCCCTCCAGCAGTACCGTACTTCTGCACAAACCGAATGCTGCTGCTGGTGGCA 713
QY 325 GlnArgAspArgAlaSerArgIlePheProHisLeuThrLeuGlnSerGlyTrpAsnAla 344
Db 714 CAGGGGAGCCAGCGCTCCCGCATCTTCCCAACTCTACCTGCGGCTCAGATGCAACCA 773
QY 345 Ala-AsnLeuGln-GlnLeuGlnAlaArgAsn-ArgValThrHisIleLeuAsnMetAla--- 362
Db 774 GCCAAACCTCGAGGAGAGCTGCCAAAGCAAGAGGTCAACCCATCTTGAACATGGGCGCC 833
QY 363 ArgGluIleAspAspPheThrProGlnArgPheThrIle-----HisAsnValArgLeu 380
Db 834 GGGAGATTGGCAACCTTCTTAACCTTGAAACGCGCTTCAACCTTACCACCAATGTGGCGCTC 893
QY 381 TrpAspGlnGlnSerAlaGlnLeuLeuProHisTrp 392
Db 894 CTGGGGAATGAGGGA-----ATCGGCCCCAGCTGG 923

RESULT 5
Bg750408 812 bp mRNA linear EST 15-MAY-2001
LOCUS 602709243F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4845707 5',
DEFINITION mRNA sequence.
ACCESSION Bg750408
VERSION Bg750408.1 GI:14061061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1685 Row: f Column: 12

```

```

FEATURES
Source
Location/Qualifiers
1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4845707"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site.1: XhoI; Site.2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"

BASE COUNT 167 a 250 c 254 g 141 t
ORIGIN

Alignment Scores:
Pred. No.: 8,76e-126 Length: 812
Score: 1301.50 Matches: 264
Percent Similarity: 97.06% Conservative: 0
Best Local Similarity: 97.06% Mismatches: 7
Query Match: 53.43% Indels: 3
DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x Bg750408 (1-812)
QY 66 GlnLysAlaProSerGlnGlnGlnLeuHisGlnAspGlnThrAspPheGlyGlnGlySer 85
Db 1 CAGAAAGCGCCGAGTACAGAGAGAGCTCCAGGGGACAGACAGACTTGGGCAAGATCC 60
QY 86 GlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeuMetValGlnLeuLeu 105
Db 61 CAGAGTCCCGCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 106 ArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArgLeu 125
Db 121 AGCGCGCAGATGACATCGCTGCGCAGCCAGCGAGGAGACCCCGCGCTCCCGCGCTC 180
QY 126 ArgTyrLeuLeuValValSerThrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 145
Db 181 CGCTACCTCGCTGAGTACTTCTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 146 LeuLeuGlnValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuValLeu 165
Db 241 CTCCGTGGCGCTGATTTCCCTGACAGCAGCTCCCGCACCTGCACCCCTGGCGCTGTTG 300
QY 166 ProLeuTrpSerAspThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSer 185
Db 301 CCCCTCTGAGTACACCCAGGTGACTTAAATGAGAGAGGCGGCTTACGCTGACGCTCT 360
QY 186 GlnGlnGlnSerArgIlePheLysProIleSerIleGlnThrMetTrpAlaThrLeuGln 205
Db 361 GGTGGCAACACCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGCCACACTCCAG 420
QY 206 ValLeuHisGlnAlaCysGlnAlaAlaLeuGlnSerGlyLeuValProGlyLysAla 225
Db 421 GATTATGACCAAGCATGTGAGAGAGCTTACAGGCGGCTTGTACCGGGTGGCAGTCC 480
QY 226 LeuThrTrpAlaSerHisTyrGlnGlnArgLeuAsnSerGlnGlnSerCysLeuAsnGlu 245
Db 481 CTCACCTGGCGCAGCCACTTACAGAGAGAGTGAATCTCGAAGAGCTGCCTCAATAG 540
QY 246 TrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGluProGlyLysSer 265
Db 541 TGGAGCGGTATAGCGCGAGCTGTGGCGCTCCACAGCCCGGAGCTGGCGGCTCC 600
QY 266 SerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpLysValLeuAspVal 285
Db 601 TCAGAACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

```


Plate: L1CM384 row: a column: 18
High quality sequence stop: 709.

FEATURES

source
1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:3689369"
/clone_lib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 172 a 252 c 266 g 156 t
ORIGIN
Alignment Scores:
Pred. No.: 1.32e-109 Length: 846
Score: 1147.50 Matches: 251
Percent Similarity: 92.96% Conservative: 0
Best Local Similarity: 92.96% Mismatches: 17
Query Match: 47.11% Indels: 8
DB: 10 Gaps: 1
US-09-761-640-4 (1-471) x BE563259 (1-846)
Oy 128 LeuLeuValValSerThrArgGluGlyGluSerGlnAspGluThrValLeuLeu 147
Db 2 CTGCTGGTAGTTCCTACACGACAGAGAGGTCTGAGCCAGATGACGGTCTCTCG 61
Oy 148 G1yValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeu 167
Db 62 GGCCTGGATTCCCTGACAGCAGCTCCGCCAGCTGACCTGGCTGGTCTTGCCTC 121
Oy 168 TTPSeraPThrGlnValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGly 187
Db 122 TGGAGTGCACCCAGATGACTTACATGAGAGCGGGGCTTACGGTGCAGCTGCTGG 181
Oy 188 GlnSerArgIlePheIlePheLysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeu 207
Db 182 CAAAGCCGATTTCAAGCCCTCTCCATCCAGACCATGSGGCCACACTCCAGGTATTG 241
Oy 208 HisGlnAlaCysGluAlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThr 227
Db 242 CACCAAGCATGTGAGAGCACTGAGCAGCGGCTTGTACGGGTGGCAGTGCCCTCACC 301
Oy 228 TrpAlaSerHisTyrGlnGlnArgLeuAsnSerGluGlnSerCysLeuAsnGluTrpThr 247
Db 302 TGGGCGAGCCTACACAGAGAGAGTCACTCCGAACAGCTGCCTCAATGAGTGGAGC 361
Oy 248 AlaMetAlaAspLeuGluSerLeuArgProProSerAlaGluProGlyGlySerSerGlu 267
Db 362 GGTATGGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCTGGCGGCTCTCAGAA 421
Oy 268 GlnGlnGlnMetGluGlnAlaAlaIleArgAlaGluLeuTrpLysValLeuAspValSerAsp 287
Db 422 CAGGAGCAGATGAGACAGCGCATCCGTGCTGAGTGTGAAGTGTGATGTCAGTGAC 481
Oy 288 LeuGluSerValThrSerLysGluIleArgGlnAlaLeuGluLeuArgLeuGlyLeuPro 307
Db 482 CTGGAGAGTGTACTTCCAAATAGATCCG-CAGGCTTGGAGCTCGCTGGGCTCCGCC 540
Oy 308 LeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnTrpAsp 327
Db 541 CTCACGACGTACCTGACTTCTACACACAG-ATGCTGCTCTGTGTGACACAGCGGGAC 599
Oy 328 ArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGluTrpAsnAlaIleAsnLeu 347
|||||

Db 600 CGAGCTCCCGCATCTTCCCCCACTACTG-GGCTCAGATGGACGACCAACCTG 658
Oy 348 G1yGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAspAsn 367
Db 659 GAGAGCTGCAGAGACAGGAGGTACCCACATCTTGACATGCC-CGGAGATGTA-AAC 716
Oy 368 PheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGluGluSerAlaGln 387
Db 717 TTCTAACCTGACCGGCTTACC-TACCACATGTGGGCTTGGGATGAGAGATCGGGCCAC 775
Oy 388 LeuLeuProHisTrpLysGluThrHisArg 397
Db 776 TGTGGCCP-----GTAGGAGACGACCGT 799
RESULT 8
LOCUS BG169297 910 bp mRNA linear EST 06-FEB-2001
DEFINITION 602321027F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4424070 5',
RNA sequence.
ACCESSION BG169297 GI:12676000
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
<http://image.llnl.gov>
Plate: L1AM1067 row: f column: 07
High quality sequence stop: 742.

FEATURES

source
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4424070"
/clone_lib="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."
BASE COUNT 189 a 270 c 304 g 146 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.49e-109 Length: 910
Score: 1147.50 Matches: 246
Percent Similarity: 92.91% Conservative: 3
Best Local Similarity: 91.79% Mismatches: 15
Query Match: 47.11% Indels: 8
DB: 12 Gaps: 2
US-09-761-640-4 (1-471) x BG169297 (1-910)
Oy 179 G1yGlyPheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGln 198
Db 2 GGGGGCTTCAAGGTGACGCTGTGTGGGCAAGCGGATTTCAAGCCATCTCCATCCAG 61
Oy 199 ThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGluAlaAlaLeuGlySerGly 218
|||||

```

Db      62 ACCATGTGGGCCACTCCAGGTATG-CACCAAGCATGTGAGGACAGCTCTAGGACGGC 120
Qy      219 LeuValProG1GlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSer 238
Db      121 CTTGTACCGGGTGGCTGCTCCTCACCCTGGCCAGCCACTACCAAGAGACTGAACCTCC 180
Qy      239 GluGlnSerCysLeuAsnGlnTrpThrAlaMetAlaAspLeuGlnSerLeuArgProPro 258
Db      181 GAACAGAGCTGCTCATATGATGACGAGCGCTATGCGGACTGGAGTCTCTGCGGCTCC 240
Qy      259 SerAlaGluProGlyGlySerSerGluGlnGluMetGlnAlaIleArgAlaGlu 278
Db      241 AGCGCCGACCTGGCGGCTCTCAGACAGAGACAGATGAGCAGCGATCCCTGCTGAG 300
Qy      279 LeuTrpValLeuAspValSerAspLeuGluSerValThrSerLysGlnIleArgGln 298
Db      301 CTGTGAAAGTGTGGATGTCACTGACCTGGAGAGTGTCTCTCCAAAGAGATCCG-CAG 359
Qy      299 AlaLeuGlnLeuArgLeuGlnLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnGln 318
Db      360 GCTCTGAGCTGGCGCTGGGGGCTCCCTCCACAGTACCGTACCTTCATCGACCAACAG 419
Qy      319 MetLeuLeuLeuValAlaGlnIleArgAspArgAlaSerArgIlePheProHisLeuTyrLeu 338
Db      420 ATGCTGCTGCTGTGGCAGACGCGGACGAGCCTCCCGCATCTTCCCAACCTCTACCTG 479
Qy      339 GlySerGluTrpAsnAlaAlaAsnLeuGlnGluGlnIleArgAsnArgValThrHisIle 358
Db      480 GGTCTGAGTGGAGCCAGCAACCTGGAGAGCTGCAGAGGAAACAGGCTACCCACATC 539
Qy      359 LeuAsnMetAlaArgGlnIleAspAsnPheTyrProGlnArgPheThrTyrHisAsnVal 378
Db      540 TTGAAATGTGGCCGGGAGATNGACAACTTCACTGAGGGCTTCACTACCAATGTCG 599
Qy      379 ArgLeuTrpAspGlnLeuSerAlaGlnLeuProHisTyrLysGlnIleArgPhe 398
Db      600 CGCTC-TGGGATGAGAGTGCGCCAGCTGCTGC--GACTGGAAGAGAGCCCGGTTT 655
Qy      399 IleGlu-AlaAlaArgAlaGlnIleTyrHisValLeuValHisCysLysMetGlyValSe 418
Db      656 ATGAGAGGCTGCAGAGACAGCAGGCAACCCAGTGTGCTGCTCACTGCAAGATGGGCTCAG 715
Qy      418 rArgSerAlaAlaThrValLeuAlaTyrAlaMetLysGln-TyrGluCysSerLeuGln 438
Db      716 CCGGTACGGCGGCACAGG--GCTGGCTATGCTCATGAGCAAGTACGAGGCGGCTGGAGC 773
Qy      438 IValAlaLeuArgHisValGln 444
Db      774 GGCT--GGGCACGTGAG 790

RESULT 9
LOCUS   BE869222 726 bp mRNA linear EST 20-OCT-2000
DEFINITION 60144503orf1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849342 5',
          mRNA sequence.
ACCESSION BE869222
VERSION   BE869222.1 GI:10317998
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
          NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE    Unpublished (1999)
JOURNAL  Contact: Robert Strausberg, Ph. D.
          Email: cga@bbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone Distribution: MGC clone distribution information can be

```

```

          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          plate: L1AM9566 row: k column: 07
          High quality sequence stop: 607.
FEATURES
          source
            1..726
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:3849342"
              /clone_lib="NIH_MGC_65"
              /tissue_type="adenocarcinoma"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: colon; Vector: pCMV-SPORT6; Site.1: NotI;
              Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.8 kb. Library constructed by Life
              Technologies."
BASE COUNT  137 a 235 c 226 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 2,11e-109 Length: 726
Score: 1144.50 Matches: 230
Percent Similarity: 96.69% Conservative: 4
Best Local Similarity: 95.04% Mismatches: 7
Query Match: 46.98% Indels: 2
DB: 12 Gaps: 1
US-09-761-640-4 (1-471) x BE869222 (1-726)
Qy      93 GluGlnArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArg 112
Db      3 GACACAGAGGAGCAGCCTGCACCTCATGTGTAACACTGCTGAGCGCCGAGATGACATCCG 62
Qy      113 LeuAlaGlnLeuGlnIleArgProArgProArgLeuArgTyrLeuValValSer 132
Db      63 CTGGCAGCCAGCTGAGGAGCACCCTCCGCTCCCGCTCCGCTACTGCTGTGATTTCT 122
Qy      133 ThrArgGlnGlyGlnLeuSerGlnAspGluThrValLeuLeuGlnIleValAspPhePro 152
Db      123 ACACGAGAGAGAGAGAGTGTGAGCCAGAGATGAGACGGTCTCTCTGGCGGTGATTTCCCT 182
Qy      153 AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrSerAspThrGln 172
Db      183 GACAGCAGCTCCCGCCAGCTGCACCTGGCGGTGCTCTCCCTCTGGAGTGCACACCCAG 242
Qy      173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlnSerArgIlePhe 192
Db      243 GTTACTTATGATGAGACAGCGGGGCTTCAGCGTACGTCTGTGGGCAAGCCGATCTTC 302
Qy      193 LysProIleSerIleGlnThrMetTrpAlaThrLeuGlnIleValLeuHisGlnAlaCysGln 212
Db      303 AAGCCCATCTCCATCCAGACCATGTGGCCACACCTCCAGTATTCACCAAGCATGTGAG 362
Qy      213 AlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyr 232
Db      363 GCAGCTCTAGGACGCGGCTGTACCGGGGTGGCAGTCCCTACCTGGGCGACCCACTAC 422
Qy      233 GlnGluArgLeuAsnSerGlnSerCysLeuAsnGlnIleTrpThrAlaMetAlaAspLeu 252
Db      423 CAGGAGAGACTGACTCCGACAGAGACTGCTCATATGATGGAGCGCTATGGCCGACCTG 482
Qy      483 GACTCTCTGGGCTCCCGACGCGCGGAGCTGGGGGTCCCTCAGAACAGAGCAGATGAG 542
Db      543 GlnAlaIleArgAlaGlnLeuTrpLysValLeuAspValSerAspLeuGlnSerValThr 292
Qy      293 SerLysGlnIleArgGlnAlaLeuGlnIleArgLysGlnIleArgProLeuGlnGlnIleTyrArg 312
Db      603 TCCAAAGAGATCCGCCAGGCTCTGGAGCTGGGCTCCCTCCCGCCAGTACCGC 662

```

QY 313 AspHeileaspangimleuleuleuValAlaInArgAspArgAlaSerArgIle 332
 Db 663 GACTTCATFCGACACAG-ATGCTGCTGTGTGGCCAGC---GGGCGAGCGCTCGCCTC 718
 QY 333 PhePro 334
 Db 719 TTCCCC 724

RESULT 10
 BE737087 919 bp mRNA linear EST 15-SEP-2000
 LOCUS 60130485F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3639229 5',
 DEFINITION mRNA sequence.
 ACCESSION BE737087
 VERSION BE737087.1 GI:10151079
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 919)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHC341 row: h column: 14
 High quality sequence start: 62
 High quality sequence stop: 769.
 Location/Qualifiers
 1. 919
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3639229"
 /clone_1b="NIH_MGC_39"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB; Site: 1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-CDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies)."

BASE COUNT 170 a 269 c 316 g 164 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.13e-103 Length: 919
 Score: 1091.50 Matches: 236
 Percent Similarity: 91.29% Conservative: 5
 Best Local Similarity: 89.39% Mismatches: 18
 Query Match: 44.81% Indels: 8
 DB: 12 Gaps: 1

US-09-761-640-4 (1-471) x BE737087 (1-919)

QY 175 leuaspGlyaspGlyGlyPheSer-ValThrSerGlyGlyInserArgIlePheLeu 194
 Db 42 TTATATGAGAGCGGGGT---TCAGGTGACGCTGTGGCAACCGCATCTTCAACC 98
 QY 194 oileserileglnhrmettrpalaThrleuGlnValleuHisGlnAlaCysGluAlaI 214
 Db 99 CATCTCATCCAGACCATGTGGCCACACTCCAGGTATG-CACCAACGATGTAGGACG 157

QY 214 aleuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrpAlaSerHisTyGlnG 234
 Db 158 TCTAGGACAGCGGCTTGTACCGGGGTGGAGTCCCTCATCGGGGACCACTACACAGA 217
 QY 234 uArgLeuAsnSerGlyGlnSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuG 254
 Db 218 GAGACTGAACCTCCGACAGACAGCTGCTCATATGAGGAGGAGGTATGGCCGACCTGAGCTC 277
 QY 254 rleuAtrProProSerAlaGluProGlyGlySerSerGlyGlnGlnGlnGlnGlnG 274
 Db 278 TCTGGGCGCTCCAGCGCGGAGCTGGCGGCTCTCAAGACAGGAGCATGAGACGGGGC 337
 QY 274 alleArgAlaGluLeuTrpLysValleuAspValSerAspLeuGlnSerValThrSer 294
 Db 338 GATCCGCTCTGAGCTGTGGAAGTGTGATGTGCTGACCTGAGAGATGTGCTTCAG 397
 QY 294 sgluileArgGlnAlaLeuGluLeuArgLeuGlyLeuProLeuGlnGlnTrpArgAsp 314
 Db 398 AGAGATCCGCGAGGCTCTGGAGCTGCGCTGGGCTCCCTCCAGCATCCGACTT 457
 QY 314 elleaspangimleuleuleuValAlaGlnArgAspArgAlaSerArgIlePhe 334
 Db 458 CATGCACACAGATGCTGTGTGTGGACAGCGGAGCGAGCTCCCATCTTCC 517
 QY 334 oHisLeuTrleuGlySerGluTrpAsnAlaAlaAsnLeuGluGlnLeuGlnArgAsn 354
 Db 518 CCACCTTACCTGGGCTGAGAGTGAACGACGCAACCTGAGAGCTCCAGAGGACAG 577
 QY 354 gvalThrHisIleLeu-AsnMetAlaArgGluIleAspAsnPheTyProGluArgPhe 374
 Db 578 GGTCCACCATCTTGTAACTATGCGCCGAGATGACAA-TTTCACCTGAGCCCTCA 636
 QY 374 hrTyHis-AsnValArgLeuTrpAspGluGlnSerAlaGlnLeuLeuProHisTrp 393
 Db 637 CTTACACAGATGTGCGCTGTGGATGAGAGTGGC-CAGGTCTCCCGCATGTGAGG 695
 QY 394 GluThrHisArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeuValHis 413
 Db 696 GAGACGACCGCTTCAATTGAGGCTCGAGAGACAGCGGACCCACCTGCTGCTCCG 755
 QY 414 -LysMetGlyValSerArgSerAlaIleThrValLeuAlaTrpAlaMetLysGln 433
 Db 756 CAAGATGGGGGCGTGCAGCGTCAGAGGGGCCCGGCTGTATGCGTGAAGACAGTAC 815
 QY 433 uCys 434
 Db 816 ATGC 819

RESULT 11
 LOCUS BM563401 943 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT_6589387 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:541645
 5', mRNA sequence.
 ACCESSION BM563401
 VERSION BM563401.1 GI:18810292
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 943)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

Plate: L1CM1916 row: e column: 06
High quality sequence stop: 618.

FEATURES

SOURCE

Location/Qualifiers
1. 943

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5441645"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 201 a 284 c 305 g 153 t
ORIGIN

Alignment Scores:

Pred. No.: 1.91e-103 Length: 943
Score: 1089.50 Matches: 236
Percent Similarity: 75.24% Conservatave: 4
Best Local Similarity: 73.98% Mismatches: 9
Query Match: 44.72% Indels: 70
DB: 13 Gaps: 3

US-09-761-640-4 (1-471) x BM563401 (1-943)

QY 25 AAlaValGlnArgGArgSerArgLeuGlnArgGlnSerPheAlaValLeuArgGlyAla 44
DB 27 GCGGTCACAGCGAAGAGTGCTGACTCCAGCGAAGCAGAGCTTGCGGTGCTCCGTGGGCT 86
QY 45 ValLeuGlyLeuGlnAspGlyGlyAspAsnAspAspAlaAlaGlnAlaSerSerGluPro 64
DB 87 GTCCCTGGGACTGTCAGAGTGAGGAGGACATGATGATGACAGAGGCCCATCTGAGCCA 146
QY 65 ThrGluValAspProSerGlnGlnGlnLeuHisGlyAspGlnThrAspPheGlyGlnGly 84
DB 147 ACAGAGAAAGCCCCAGAGTACAGAGAGAGCTCCAGGGGACCAACAGACTTCGGGCAAGGA 206
QY 85 SerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeuMetValGlnLeu 104
DB 207 TCCCAAGATCCCAAGACAGAGAGAGAGAGAGACCTGACCTCATGTGACAGCTG 266
QY 105 LeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuAlaProArgProProArg 124
DB 267 CTGAGGCGCAGAGATGACATCCGCTGACGCCACGCTGGAGGACCCCGGCTCCCGCG 326
QY 125 LeuArgTyrLeuLeuValValSerThrArgGlnGlnGlnGlnGlnSerGlnAspGluThr 144
DB 327 CTGCGCTACCTGCTGTTGTTCTACACAGAGAGAGAGAGAGTCTGAGCCAGATGAGAGC 386
QY 145 ValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThrLeuGlyLeuVal 164
DB 387 GTCCTCTCTGGGGGTGATTTCCCTGACAGACAGCTCCCGACCTGAGACCTGGGCTGGTC 446
QY 165 LeuProlLeuTrpSerAspPheGlnValTyrLeuAspGlyAspGlyGlyPheSerValThr 184
DB 447 TTGCCCCCTCTGGAGTACACCCAGAGTGAATTGATGAGAGGGGGGCTTCAGCGTGGAGC 506
QY 185 SerGlyGlyGlnSerArgIlePheLysProLieserIleGlnThrMetTrpAlaThrLeu 204
DB 507 TCTGTTGGGCAAAACCCGATCTTCAAGCCCATCTCCATCCAGCAATGTGG----- 557
QY 205 GlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySer 224
DB 557 ----- 557
QY 225 AlaLeuThrTrpAlaSerHisTyrGlnGlnGlnArgLeuAsnSerGlnGlnSerCysLeuAsn 244

DB 557 ----- 557
QY 245 GluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGlnProGlyGly 264
DB 557 ----- 557
QY 265 SerSerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrpLysValLeuAsp 284
DB 558 TCTCTCAGAACAGAGACAGATGAGAGAGAGAGAGTCCGTGCTGAGACTGTGGAATGTGGAT 617
QY 285 ValSerAspLeuGlnSerValThrSerLysGluIleArgGlnAlaLeuGlnLeuArgLeu 304
DB 618 GTACAGACTCTGAGAGAGTGTCACTTCCCAAGAGATCCCGCAGAGCTGTGAGCTGGCGCTG 677
QY 305 GlyLeu-ProLeuGlnGlnTyrArg-AspPheIleAsp-AsnGlnMetLeuLeuLeu--- 322
DB 678 GGGCTCCCCCTCCAGACATACCGGAGCTTATGAAACAAATGGCTGCTGCTGG 737
QY 323 ValAlaGln-----ArgAspArgAlaSerArgIlePheProHisLeu 336
DB 738 GTGGCAGACCGCGGGAGCCGAGCCCTCCCGGCATCTTCCCGCCACCTT 788

RESULT 12

RG340220
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1258 row: p column: 02
High quality sequence stop: 630.
Location/Qualifiers

FEATURES

SOURCE

1. 932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4556401"
/clone_lib="NIH_MGC_46"
/tissue_type="Jelomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

ORIGIN

160 a 284 c 326 g 162 t

Alignment Scores:

Pred. No.: 5.56e-103 Length: 932
Score: 1085.00 Matches: 236
Percent Similarity: 90.98% Conservatave: 6

Best Local Similarity:	88.72%	Mismatches:	21
Query Match:	44.54%	Indels:	7
	12	Gaps:	1
US-09-761-640-4 (1-471) x BG340220 (1-932)			
QY	1	MetaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly	20
Db	43	ATGGCCCTGGTACAGTAGCGGTTCCGCCCGGGCAGCGGGCTCCACAGCCGCTGGGG	102
QY	21	ProTrpAspGlnAlaValGlnArgSerArgLeuGlnAlaArgGlnSerPheAlaVal	40
Db	103	CCCTGGGACCGGGGCTCCAGGAGGAGTGCATCCAGCGAGGACAGAGCTTTCGGGTG	162
QY	41	LeuArgGlyAlaValLeuGlyLeuGlnAspGlyValAspAsnAspAlaIaGluAla	60
Db	163	CTCCGTGGGGCTGCTCGTGGAGTGCAGATGAGGGGCAATGATGATGACACAGAGCC	222
QY	61	SerSerGluProThrGluLysAlaProSerGluGluLeuHisGlyAspGlnThrAsp	80
Db	223	AGTTCTGAGCCACAGAGAGAGCCCGGAGTGAAGAGAGCTCCAGGGGACACAGACAGC	282
QY	81	PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnAlaGlnHisLeuHisLeu	100
Db	283	TTCGGGCAAGGATCCACAGATCCCGAGAGCAGAGGAGGAGCAGACCTCCACCTC	342
QY	101	MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaIaGlnLeuGluAlaPro	120
Db	343	ATGGTACAGCTGCTGAGCCGAGGATGACATCCGCTGGCAGCCAGCTGGAGGACACC	402
QY	121	ArgProArgLeuArgTyrLeuLeuValValSerThrArgGluGlyLysLeuSer	140
Db	403	CGGCT-CCCGCGCTCGCTACCTGCTGAGTTTCTACAGAGAGAGAGAGGAGTGAAGC	461
QY	141	GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerProSerCysThr	160
Db	462	CAGGTGAGAGCGTCTCTCGGCGTGGATTCTCCAGACAGCTCCCGCGCTGCACC	521
QY	161	LeuGlyLeuValLeuProLeuThrSerAspPheGlnValTyrLeuAspGlyAspGly	180
Db	522	CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	581
QY	181	PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet	200
Db	582	TTCAGCGTGAAGTGTGGGGGCAAGCCGAGTCTCAAGCCATC-TCCATCAAGACATG	640
QY	201	TrpAlaThrLeuGln-ValLeuHisGlnAlaCysGluAlaIaLeuGlySerGlyLeu-V	220
Db	641	TGGGCGACACTCCAGGCTATGGCACCCAGCATGTGAGCAGCTTAAGCGCGGCTTTG	700
QY	220	AlProGlyLysSerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGlu	240
Db	701	TTCCTGTTGGGGAGTCCCTCAAGCTGGGAGCACT-TCCCGGGGGAGGAGTCACTCGA-C	758
QY	240	InsSerCysLeuAsnGluTrpThrAlaMetAlaAspLeuGlnSerLeuArgProProSerA	260
Db	759	AGAGTGCTTCATGAGAGTGGCGGGGTATGCCAGACCTGGAGGTTCTGGGG---CTCCACG	815
QY	260	IaGluProGlyGly 264	
Db	816	CCGACGCTGGGGC 829	
RESULT 13			
LOCUS	BE873337	897 bp	mRNA linear EST 20-OCT-2000
DEFINITION	601450290F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3854158 5',		
ACCESSION	BE873337		mRNA sequence.
VERSION	BE873337.1	GI:10322113	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia: Eutheria: Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 897)		
TITLE	NIH-MGC http://mgi.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: L1AM9579 row: c column: 23 High quality sequence stop: 632.		
FEATURES	Location/Qualifiers		
source	1. 897 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3854158" /clone_lib="NIH-MGC_65" /tissue_type="adenocarcinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."		
BASE COUNT	168 a	254 c	322 g 152 t 1 others
ORIGIN			
Alignment Scores:			
Pred. No.:	1.49e-98	Length:	897
Score:	1042.50	Matches:	222
Percent Similarity:	95.28%	Conservative:	0
Best Local Similarity:	95.28%	Mismatches:	10
Query Match:	42.80%	Indels:	5
DB:	12	Gaps:	1
US-09-761-640-4 (1-471) x BE873337 (1-897)			
QY	53	AspAsnAspAspAlaIaGluAlaSerSerGluProThrGluLysAlaProSerGluGln	72
Db	2	GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	61
QY	73	GluLeuHisGlyAspGlnThrAspPheGlyGlnGlySerGlnSerProGlnLysGlnGlu	92
Db	62	GAGCTCCAGGGGAGACAGACACTTCGGGCAAGGATCCAGAGTCCCGCAGAGGAGG	121
QY	93	GluGlnArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArg	112
Db	122	GAGCAGAGGACACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG	181
QY	113	LeuAlaIaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuLeuValValSer	132
Db	182	CTGGCAGCCACCTGAGAGCACCCGGCTCCCGGCTCCGCTACCTGCTGATGATTTCT	241
QY	133	ThrArgGluGlyGlnGlyLeuSerGlnAspGluThrValLeuLeuGlyValAspPhePro	152
Db	242	ACACGAGAGGAGAGGATGAGCAGGATGAGAGGCTCTGCTGGGCTGGGATTTCTCT	301
QY	153	AspSerSerSerProSerCysThrLeuGlyLeuValLeuProLeuThrSerAspPheGln	172
Db	302	GACAGCAGCTCCCGCAGCTGACACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	361
QY	173	ValTyrLeuAspGlyAspGlyLysPheSerValThrSerGlyGlnSerArgIlePhe	192
Db	362	GCTACTTATGATGAGAGAGGGGCTTACCGCTGAGCTGTGGGCAAGCGGATCTTC	421
QY	193	LysProIleSerIleGlnThrMetTrpAlaThrLeuGlnValLeuHisGlnAlaCysGlu	212
Db	422	AAGCCATCTCCATCCAGACCATGTGGGGCCACATCCAGATGATTCACCAAGCATGTGAG	481


```

/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT      138 a      213 c      251 g      114 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3.7e-93      Length:      716
Score:          989.50      Matches:      207
Percent Similarity: 97.64%      Conservative: 0
Best Local Similarity: 97.64%      Mismatches: 5
Query Match:      40.62%      Indels:      4
DB:              12      Gaps:      0

```

US-09-761-640-4 (1-471) x BE907514 (1-716)

```

OY      1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
      |||||||
DB      77 ATGGCCCTGTCACAGTAGTCCGTTGCCCCGGGAGCGGGCCCTCCACGCCCGTGGGG 136
      |||||||

OY      21 ProTTPASPGLnAlaValGlnArgArgSerArgLeuGlnArgGlnSerPheAlaVal 40
      |||||||
DB      137 CCTGGGACCGGCGGTCAGGAGGAGTCCAGCTCCAGCGAAGGACAGACTTTGGCGTG 196
      |||||||

OY      41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAspAlaAlaGluAla 60
      |||||||
DB      197 CTCCTGGGGGCTGCTCTGGGATGCGAGATGAGGGGACATGATGATGATCAGCAGAGGCC 256
      |||||||

OY      61 SerSerGluProThrGluLysAlaProSerGluGluGluLeuHisGlyAspGlnThrAsp 80
      |||||||
DB      257 AGTTCTGAGCCACAGAGAGAGAGCCCGAGTGAAGAGAGCTCCACGGGAGCCAGACAGAC 316
      |||||||

OY      81 PheGlyGlnGlySerGlnSerProGlnLysGlnGluGlnArgGlnHisLeuHisLeu 100
      |||||||
DB      317 TTCGGGCAAGATCCAGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
      |||||||

OY      101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaAlaGlnLeuGluAlaPro 120
      |||||||
DB      377 ATGTACACAGCTCTAGAGCGCGAGATGACATCCCGCTGGCAGCCAGCTGGAGGACCCC 436
      |||||||

OY      121 ArgProProArgLeuArgGlyLeuLeuValValSerThrArgGluGlyLeuSer 140
      |||||||
DB      437 CGGCTCCCGGCTCCGCTACCTGCTGAGTTCTACACGAGAGAGAGAGAGGTTGAGCC 496
      |||||||

OY      141 GlnAspGluThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
      |||||||
DB      497 CAGGATGAGACGCTCTCTGGGCGTGGATTTCCCTGACAGCAGCTCCCGCAGCTGCACC 556
      |||||||

OY      161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValThrLeuAspGlyAspGlyGly 180
      |||||||
DB      557 CTGGGCTGCTGCTTGGCCCTCTGGAGTGACACCCAGGTACTTAGATGAGAGCGGGGSC 616
      |||||||

OY      181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
      |||||||
DB      617 TTCAGCGTAGCTGAGGGGCAAG-CGGATCTCAAGCCATC-TCCATCCAGAC-ATG 673
      |||||||

OY      201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGln 212
      |||||||
DB      674 TGGGGACACTCCAGGTATTGGC-CAAGCATGTGAA 708
      |||||||

```

Search completed: April 12, 2003, 10:52:46
Job time : 3880 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: April 11, 2003, 06:56:50 ; Search time 344 Seconds
(without alignments)
1201.006 Million cell updates/sec

Title: US-09-761-640-4
2436
Sequence: 1 MATVTSRSPGSGASTPVG.....PNEGFLROLIYQGLTART 471

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 593429 seqs, 438583890 residues
Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n model -DEV=xlh
-O=/cgn2.1/USFO_spool/US09761640/unat_08042003.141440.20481/app.query.fasta.1.647
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=trnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=ext -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09761640.ecgn.1.1.99.ecgnat.08042003.141440.20481
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Published.Applications_NA.*
1: /cgn2.6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2.6/ptodata/2/pubna/PC1_NEM_PUB.seq.*
3: /cgn2.6/ptodata/2/pubna/US06_NEM_PUB.seq.*
4: /cgn2.6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2.6/ptodata/2/pubna/US07_NEM_PUB.seq.*
6: /cgn2.6/ptodata/2/pubna/PC1US_PUBCOMB.seq.*
7: /cgn2.6/ptodata/2/pubna/US08_NEM_PUB.seq.*
8: /cgn2.6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2.6/ptodata/2/pubna/US09_NEM_PUB.seq.*
10: /cgn2.6/ptodata/2/pubna/US09_PUBCOMB.seq.*
11: /cgn2.6/ptodata/2/pubna/US10_NEM_PUB.seq.*
12: /cgn2.6/ptodata/2/pubna/US10_PUBCOMB.seq.*
13: /cgn2.6/ptodata/2/pubna/US60_NEM_PUB.seq.*
14: /cgn2.6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	100.0	1416	9	US-09-955-732-20
2	2436	100.0	2704	10	US-09-761-640-1
3	2436	100.0	2704	10	US-09-761-640-7
4	2426	99.6	1980	9	US-09-955-732-1

5	2426	99.6	2852	10	US-09-761-640-2	Sequence 2, Appl1
6	2061.5	84.6	2540	10	US-09-761-640-3	Sequence 3, Appl1
7	1045	42.9	1711	10	US-09-775-925-5	Sequence 5, Appl1
8	980	40.2	1949	10	US-09-775-925-1	Sequence 1, Appl1
9	589	24.2	1052	10	US-09-775-925-7	Sequence 7, Appl1
10	539	22.1	969	10	US-09-808-701-12	Sequence 12, Appl1
11	256	10.5	654	10	US-09-815-419-3	Sequence 3, Appl1
12	256	10.5	1292	10	US-09-815-419-1	Sequence 1, Appl1
13	235.5	9.7	1337	9	US-09-964-899-46	Sequence 46, Appl1
14	230	9.4	1391	10	US-09-878-574-1411	Sequence 1411, Ap
15	227.5	9.3	1998	10	US-09-816-494-3	Sequence 3, Appl1
16	227.5	9.3	2470	10	US-09-964-824A-555	Sequence 355, Ap
17	227.5	9.3	2470	10	US-09-880-107-3310	Sequence 3310, Ap
18	227.5	9.3	3496	10	US-09-964-277-1	Sequence 1, Appl1
19	227.5	9.3	3544	10	US-09-816-494-1	Sequence 1, Appl1
20	227	9.3	2436	10	US-09-917-800A-485	Sequence 485, Ap
21	226.5	9.3	2139	9	US-09-865-993-3	Sequence 3, Appl1
22	226.5	9.3	2479	9	US-09-865-993-10	Sequence 10, Appl1
23	222.5	9.1	705	10	US-09-963-204-3	Sequence 3, Appl1
24	222.5	9.1	867	10	US-09-963-204-13	Sequence 13, Appl1
25	222.5	9.1	1160	10	US-09-963-204-1	Sequence 1, Appl1
26	216.5	8.9	3059	10	US-09-770-595A-2	Sequence 2, Appl1
27	216.5	8.9	6356	10	US-09-770-595A-1	Sequence 1, Appl1
28	215.5	8.8	2000	10	US-09-969-708-276	Sequence 276, Ap
29	215.5	8.8	2000	10	US-09-919-497-10	Sequence 10, Appl1
30	215.5	8.8	2000	10	US-09-880-107-3774	Sequence 3774, Ap
31	215.5	8.8	2303	10	US-09-880-107-3849	Sequence 3849, Ap
32	215	8.8	2064	9	US-09-736-457-825	Sequence 825, Ap
33	215	8.8	2064	9	US-09-902-941-825	Sequence 825, Ap
34	215	8.8	2064	9	US-09-849-626-825	Sequence 825, Ap
35	215	8.8	2064	9	US-10-017-754-825	Sequence 825, Ap
36	215	8.8	2109	9	US-09-736-457-826	Sequence 826, Ap
37	215	8.8	2109	9	US-09-902-941-826	Sequence 826, Ap
38	215	8.8	2109	9	US-09-849-626-826	Sequence 826, Ap
39	215	8.8	2109	9	US-10-017-754-826	Sequence 826, Ap
40	215	8.8	2444	9	US-10-071-768-29	Sequence 29, Appl1
41	214	8.8	1447	10	US-09-738-885-1	Sequence 1, Appl1
42	214	8.8	1691	9	US-09-860-670-19	Sequence 19, Appl1
43	209.5	8.6	2453	12	US-10-005-858-1	Sequence 1, Appl1
44	209.5	8.6	3332	10	US-09-964-277-20	Sequence 20, Appl1
45	207	8.5	1238	9	US-09-736-457-803	Sequence 803, Ap

ALIGNMENTS

RESULT 1
US-09-955-732-20
; Sequence 20, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-955-732-20

Alignment Scores:
Pred. No.: 2,64e-258
Score: 2436.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9
Gaps: 0
Length: 1416
Matches: 471
Conservative: 0
Indels: 0

US-09-761-640-4 (1-471) x US-09-955-732-20 (1-1416)

```
OY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
  1 ATGGCCCTGGGTACAGAGAGCGCTTCGCCCGGGGAGCGGGCGCTCCACCCCGGTGGG 60
OY 21 ProTrrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
  61 CCCGGGACACAGGGGCTCCAGCGAAGAGAGTCCAGCTCCAGCAAGGACAGAGCTTGGGTTG 120
OY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAspAsnAspAspAlaGlnAla 60
  121 CTCGGTGGGCTGTCTGGGAGTCTCAGAGATGAGAGGAGGACATGATGACAGCAGAGGCC 180
OY 61 SerSerGluProThrGlnLysAlaProSerGlnGlnGlnLeuHisGlyAspGlnThrAsp 80
  181 AGTTCTGAGCAACAGAGAGAGGCCCGCAGTGAAGAGAGGCTCCAGGGGAGCACACAGAC 240
OY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeu 100
  241 TTCGGGCAAGGATCCAGAGTCCCGACAGACAGAGAGAGAGGACAGGACCTGCACCTC 300
OY 101 MetValGlnLeuLeuLeuArgProGlnAspAspIleArgLeuAlaGlnLeuAlaPro 120
  301 ATGGTACAGCTGCTGAGGCCGACAGATGACATCCGCTGGCAGCCAGCTGAGAGCACCC 360
OY 121 ArgProBrrArgLeuArgTyrLeuLeuValValSerThrArgGlnGlyGlnGlyLeuSer 140
  361 CGGCTGCCCGGGCTCCGCTACCTGCTGTAGTTCTACAGAGAGAGAGAGAGGCTGAGC 420
OY 141 GlnAspGlnThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
  421 CAGATTAAGACGGTCTCTGGGCGTGGATTTCCCTACAGCAGCTCCCGCAGCTGCACC 480
OY 161 LeuGlyLeuValLeuProLeuTrrPheSerAspThrGlnValTyrLeuAspGlyAspGly 180
  481 CTGGGCTGGTCTTGGCCCTCGAGTGAACACCCAGCTGACTAGTGAAGAGAGGGGGC 540
OY 181 PheSerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMet 200
  541 TTCAGGCTGACGCTGCTGGTGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
OY 201 TrpAlaThrLeuGlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuVal 220
  601 TGGGCCACACTCCAGGATATGACACCAAGCATGTGAGGACCTTAGGCAACGGGCTTGTGA 660
OY 221 ProGlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGlnLeuArgLeuAsnSerGln 240
  661 CCGGGTGGAGTGGCCCTCAGCTGGGCCAGCACTACAGAGAGAGACTGGAAGCTCCGAACAG 720
OY 241 SerCysLeuAsnGlnTrrPheAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
  721 AGCTGCTCATATGATGAGGAGGCTATGGCCGAGCTGAGTCTGCGGGCTCCAGCGCC 780
OY 261 GluProGlyLysSerSerGlnGlnGlnGlnMetGlnGlnAlaIleArgAlaGlnLeuTrr 280
  781 GAGCCCTGGGGGTCTCTAGACAGAGAGACAGATGAGAGCGATCCGTGTGAGCTGTGG 840
OY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerLysGlnIleArgGlnAlaLeu 300
  841 AAGGTGTGATGTCACTGAGTGAAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
OY 301 GluLeuArgLeuGlnLeuProLeuGlnGlnTrrArgAspPheIleAspAsnGlnMetLeu 320
  901 GAGCTCGGCTGGGGCTCCCGCTCCAGACAGTACCGTACTTCATCGACCAACAGATGCTG 960
OY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTrrLeuGlySer 340
  961 CTGCTGGTGGCAAGGGGAGCGAGGCTCCCGCATCTTCCCGACCTTCACTCGGGCTCA 1020
OY 341 GluTrrAsnAlaAlaAsnLeuGlnGlnGlnLeuGlnArgAsnArgValThrHisIleLeuAsn 360
  1021 GAGTGAACCGACCAACCTGAGAGAGCTGACAGAGACAGGGGTCAACCCACATCTTGAAC 1080
```

```
OY 361 MetAlaArgGlnIleAspAsnPheTyrProGlnArgPheThrTrrHisAsnValArgLeu 380
  1081 ATGGCCCGGAGAGATTGACAACTTACCCTGAGCGCTTCACTTCAACCAATGTGGCCTC 1140
OY 381 TrpAspGlnGlnSerAlaGlnLeuLeuProHisTrrLysGlnTrrHisArgPheIleGln 400
  1141 TGGGATGAGAGTGGGCCCGAGCTCTCCGCACTGGAAGAGAGACGACCGCTTCATTGAG 1200
OY 401 AlaAlaArgAlaGlnGlnTrrHisValLeuValHisCysLysMetGlyValSerArgSer 420
  1201 GCTGCAAGAGCAGAGGCGACCCACGCTGCTGTCCATGCAAGATGAGGCTGCAGCGCTCA 1260
OY 421 AlaAlaThrValLeuAlaTrrAlaMetLysGlnTrrGlySerLeuGlnGlnAlaLeu 440
  1261 GCGGCCACAGTGTGGCTTATGCCATGACAGACAGTACGAATGCAACCTGAGAGAGCCCTG 1320
OY 441 ArgHisValGlnGlnLeuLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
  1321 CGCCAGCTGACAGAGCTCCGCGCCATGCGCCCGCCCAACCTGGCTTCTGCGCACCTG 1380
OY 461 GlnIleTyrGlnGlyIleLeuThrAlaArgThr 471
  1381 CAGATCTACCAAGGAGCATCTGACGGCCAGAAC 1413
```

RESULT 2

```
US-09-761-640-1
; Sequence 1, Application US/09761640
; Patent No. US20020137042A1
;
; GENERAL INFORMATION:
; APPLICANT: MET, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: C1000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-1
```

Alignment Scores:

```
Pred. No.: 6,61e-258 Length: 2704
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
```

US-09-761-640-4 (1-471) x US-09-761-640-1 (1-2704)

```
OY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
  1 ATGGCCCTGGGTACAGAGAGCGCTTCGCCCGGGGAGCGGGCGCTCCACCCCGGTGGG 153
OY 21 ProTrrAspGlnAlaValGlnArgArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
  61 CCCGGGACACAGGGGCTCCAGCGAAGAGAGTCCAGCTCCAGCAAGGACAGAGCTTGGGTTG 213
OY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAspAsnAspAspAlaGlnAla 60
  121 CTCGGTGGGCTGTCTGGGAGTCTCAGAGATGAGAGGAGGACATGATGATGACAGAGGCC 273
OY 61 SerSerGluProThrGlnLysAlaProSerGlnGlnGlnLeuHisGlyAspGlnThrAsp 80
  181 AGTTCTGAGCAACAGAGAGAGGCCCGCAGTGAAGAGAGGCTCCAGGGGAGCACACAGAC 333
OY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeu 100
  241 TTCGGGCAAGGATCCAGAGTCCCGACAGACAGAGAGAGAGGACAGGACCTGCACCTC 393
```

QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaIleGlnLeuGlnAlaPro 120
Db 394 ATGGTACAGCTGTGTGAGCGCCAGAGTACATCCGCTGTGGACGCCACCTGGAGCCACCC 453
QY 121 ArgProProAlaGlyLeuArgTyrLeuLeuValValSerThrArgGlyGlyGlyLeuSer 140
Db 454 CGGCTCCCGCCGCTCCGCTACCTGCTGTAGTTCTTACACGAGAGGAGGAGTGTGAGC 513
QY 141 GlnAspGlnThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Db 514 CAGGATGAGACGGTCTCTGTGGCTGGATTTCTCTGACAGCAGCTCCGCCAGCTGACAC 573
QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
Db 574 CTGGCGCTGGCTTCTCCCTCGGAGTACACCCAGGCTACTTGAATGAGAGACGGGGC 633
QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerTleGlnThrMet 200
Db 634 TTCACGCTGACGCTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 693
QY 201 ThrAlaThrLeuGlnValLeuGlnIleGlnAlaGlyValAlaLeuGlySerGlyLeuVal 220
Db 694 TGGGCACTCCAGCTCAGATTTGACCAACAGATGTAGGACGCTTAGGCGCGCTTGTGA 753
QY 221 ProGlyGlySerAlaLeuThrThrAlaSerHisTyrGlnGlnArgLeuAsnSerGlnGln 240
Db 754 CCGGTGGACAGTCCCTCCTACCTGAGCCAGCCATCCAGAGAGAGCTAGCTCCGACAC 813
QY 241 SerCysLeuAsnGlnTyrThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAla 260
Db 814 AGCTCCCTCAATGATGAGTGGAGCGCTATGGCCGACCTGAGTCTCTCGGCTCCAGCCGC 873
QY 261 GlnProGlyGlySerSerGlnGlnGlnIleMetGlnGlnAlaIleArgAlaGlnLeuThr 280
Db 874 GAGCTGGCGGGTCTCTCAACACAGAGAGATGAGAGCGGCGATCCGTGTAGTGTGG 933
QY 281 LysValLeuAspValSerAspLeuGlnSerValThrSerLysGlnIleArgGlnAlaLeu 300
Db 934 AAGGTGTGATGTCAGTACGACCTGGAGAGTGTCACTTCCAAAGATCCGCCAGGCTCTG 993
QY 301 GlnLeuArgLeuGlyLeuProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeu 320
Db 994 GAGCTGCCCTGGGGCTCCCTCCAGCAGTACCGTGAATTCATGACAAACAGAGATGCTG 1053
QY 321 LeuLeuValAlaGlnArgAspArgAlaSerArgIlePheProHisLeuTyrLeuLysSer 340
Db 1054 CTGCTGTGGCCACAGCGGAGCCGAGCTCCGCTCCGCTCCCTGCTACCTGGGCTCA 1113
QY 341 GlnThrAsnAlaAlaAsnLeuGlnGlnLeuGlnArgAsnArgValThrHisIleLeuAsn 360
Db 1114 GAGTGGAAAGCAGCAACCTGGAGAGCTGCGAGAGAACAGGGTCAACCCACACTTTGAAC 1173
QY 361 MetAlaArgGlnIleAspAsnPheTyrProGlnArgPheThrTyrHisAsnValArgLeu 380
Db 1174 ATGGCCCGGAGATTGACAACTTTCACCTGAGCCCTTCACTACCAAAATGTGGCCTC 1233
QY 381 TrpAspGlnGlnSerAlaGlnLeuLeuProHisTyrLysGlnThrHisArgPheIleGln 400
Db 1234 TGGGATGAGAGTCCGCGCCAGCTGCTGGCGCAGTGAAGAGACGACGCGCTTCAATTGAG 1293
QY 401 AlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSer 420
Db 1294 GCTGCAAGAGCAGAGGAGCAGCAGTGTGTCACCTGCAAGATGGGGTGCAGCGGCTCA 1353
QY 421 AlaAlaThrValLeuAlaTyrAlaMetLysGlnTyrLysSerLeuGlnAlaLeu 440
Db 1354 GCGGCAAGAGTCTGTGCTATGCTGCAAGAGTACGAATGACAGCTGAGAGAGCCCTG 1413
QY 441 ArgHisValGlnGlnLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
Db 1414 CGCCACGTGACGAGACTCCGCGCCATCGCCGCCCAACCTGTGGCTTCTGTGGCAGCTG 1473

QY 461 GlnIleTyrGlnGlyTleLeuThrAlaArgThr 471
Db 1474 CAGATCTACAGGCGCATCTCTGAGCGCCAGAAC 1506

RESULT 3
US-09-761-640-7
; Sequence 7, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: C1000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-7

Alignment Scores:
Pred. No.: 6,61e-258 Length: 2704
Score: 2436.00 Matches: 471
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-761-640-4 (1-471) x US-09-761-640-7 (1-2704)

QY 1 MetAlaLeuValThrValSerArgSerProProGlySerGlyAlaSerThrProValGly 20
Db 94 ATGGCCCTGTGTACAGTACAGTACAGCTTCCGCCCCGCGGACAGCGGCTCCACGCGCTGGGG 153
QY 21 ProTrpAspGlnAlaValGlnArgSerArgLeuGlnArgArgGlnSerPheAlaVal 40
Db 154 CCTGGGACACGCGGCTCCAGGAGAGAGTGCAGCTCCAGCAGGAGCAAGCAAGCTTTCGGGTG 213
QY 41 LeuArgGlyAlaValLeuGlyLeuGlnAspGlyLysAspAsnAspAlaIleGlnAla 60
Db 214 CTCCGTGGGGGCTGCTCCTGGAGTCCAGATGAGAGGAGCAATGATGATGACAGAGGCC 273
QY 61 SerSerGlnProThrGlnLysAlaProSerGlnGlnGlnLeuHisGlyAspGlnThrAsp 80
Db 274 AGTTCTGAGCCCAAGAGAGAGGCCCCGAGTGAAGAGAGCTCCACGCGGAGCAGAC 333
QY 81 PheGlyGlnGlySerGlnSerProGlnLysGlnGlnGlnArgGlnHisLeuHisLeu 100
Db 334 TTTCGGGCAAGGATCCCAAGTCCCAAGACAGAGAGAGAGAGAGCAGCAGCTGCACCTC 393
QY 101 MetValGlnLeuLeuArgProGlnAspAspIleArgLeuAlaIleGlnLeuGlnAlaPro 120
Db 394 ATGGTACAGCTGTGTGAGCGCCAGAGTACATCCGCTGTGGACGCCACCTGGAGCCACCC 453
QY 121 ArgProProAlaGlyLeuArgTyrLeuLeuValValSerThrArgGlyGlyGlyLeuSer 140
Db 454 CGGCTCCCGCCGCTCCGCTACCTGCTGTAGTTCTTACACGAGAGGAGGAGGAGTGTGAGC 513
QY 141 GlnAspGlnThrValLeuLeuGlyValAspPheProAspSerSerSerProSerCysThr 160
Db 514 CAGGATGAGACGGTCTCTGTGGCTGGATTTCTCTGACAGCAGCTCCGCCAGCTGACAC 573
QY 161 LeuGlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyGly 180
Db 574 CTGGCGCTGGCTTCTCCCTCGGAGTACACCCAGGCTACTTGAATGAGAGACGGGGC 633
QY 181 PheSerValThrSerGlyGlyGlnSerArgIlePheLysProIleSerTleGlnThrMet 200
Db 634 TTCACGCTGACGCTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 693


```
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775.925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-775-925-5

Alignment Scores:
Pred. No.: 2.87e-105 Length: 1711
Score: 1045.00 Matches: 200
Percent Similarity: 66.37% Conservative: 94
Best Local Similarity: 45.15% Mismatches: 107
Query Match: 42.90% Indels: 42
DB: 10 Gaps: 4

US-09-761-640-4 (1-471) x US-09-775-925-5 (1-1711)
QY 36 GlnserPheAlaValleuArgGlyAlaValleuGlyLeuGlnAspGlyGlyAspAsnasp 55
   :::::::::: :::::::::: ::::::::::
DB 240 GAGACCTTCTAACTGCTCAAGAGGTCCTCTTTCTTCAACAGC----- 284
QY 56 AspaAlaAlaValleuSerSerGluProThrGluysAlaProSerGluGluGluLeuHis 75
DB 284 ----- 284
QY 76 GlyaspGlnThrAspPheGlyGlnGlySerGlnSerProGlnIlySglnGluGlnIn--- 94
DB 285 -----GGAATGCTCATCCACCAACCAATCAAGCAGCAGAGAGGAGAC 326
QY 95 -----ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGln 108
DB 327 AAGCATGCGAGCGATCTCCACAGCATCTCCACAGCAATGTTCTATTTCTCCGCCAGAA 386
QY 109 AspAspIleArgLeuAlaAlaGlnLeuGlnAlaProArgProProArgLeuArgGlyLeu 128
DB 387 GACACATCATGAGCTGGCTGAAGACTGGAAGTACTTACCAAGATCAACAGCTATATG 446
QY 129 LeuValleuSerThrArgGluGlyGlnGlyLeuSerGlnAspGlnThrValleuLeuGly 148
DB 447 GTAGTGGTTCAACTAATGTAGACAA-----GACACTGAAGAAAGCATGCTCTAGGA 500
QY 149 ValAspPheProAspSerSerSerProSerGlyThrLeuGlyLeuValleuProLeuTrp 168
DB 501 ATGATTTCTCTCTCTCAATGACAGTAGTACCTGTACATGGGCTTATGCTTCTCTCTG 560
QY 169 SerAspThrGlnValTyrLeuAspGlyLeuGlyPheSerValThrSerGlyGln 188
DB 561 AGCAGACAGCTAAATTCATTTGATGTGATGTGGTTCAGATGATGACAGGATTAACAGA 620
QY 189 SerArgIlePheIlySerProIleSerIleGlnThrMetTrpAlaThrLeuGlnValleuHis 208
DB 621 GTTCACATATTTCAACCTGTATCTGTGCGCAAGTGTGCTGCAACAGCTTATAC 680
QY 209 GlnAlaGlyGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySerAlaLeuThrTrp 228
DB 681 AAGCTTTGTGAAGTCCGACAGCCGATACACTACACAGGAGCTTATTTCTCTCTG 740
QY 229 AlaSerHisTyrGlnGluArgLeuAsnSerGluGlnSerCysLeuAsnGluTrpThrAla 248
DB 741 GTGAGTTATATAGAGAGCATATCACTACATCAATCTCAGTCAATGATGAAATGCA 800
QY 249 MetaAlaAspLeuGlnSerLeuArgPro-----ProSerAlaGluProGlyGlySerSer 266
DB 801 ATGCAAGATGTACAGTCCACCGCCGACTCTCCACGCTCTCTTCCAGCATACCTACT 860
```

```
QY 267 GlnGlnGlnMetGlnAlaIleArgAlaGluLeuTrpIlySValleuAspValSer 286
   :::::::::: :::::::::: ::::::::::
DB 861 GACCTGAGAACGACAGAAAGCCTATTTAAACCAATTTAAGGAGATCATATGATCAGAG 920
QY 287 AspLeuGlnSerValThrSerIlySglnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeu 306
DB 921 GATTTGGAGAAATATATACATCCAAAGAGATAGAACAGATGGAATGGAATGCTGTGC 980
QY 307 ProLeuGlnGlnTyrArgAspPheIleAspAsnGlnMetLeuLeuValAlaGlnArg 326
DB 981 AACTTGGCGGAATTCAGAAATTTATAGCAATGAAATGATGATCTTGGTCAATATG 1040
QY 327 AspArgAlaSerArgIlePheProHisLeuTyrLeuGlySerGlyLeuTrpAspAlaAlaAsn 346
DB 1041 GATAGCCCTACACAGATATTTGAGCATGTCTTCCTGGGCTCAGATGGAATGCCCTCCAC 1100
QY 347 LeuGlnGluLeuGlnArgAsnArgValThrHisIleLeuAsnMetAlaArgGluIleAsp 366
DB 1101 TTAGAGGACTTACAGAACCGAGGGGTACGATATATCTTGAATGTCTACAGAGATAGAT 1160
QY 367 AsnPheTyrProGluArgPheThrTyrHisAsnValArgLeuTrpAspGluGlnSerAla 386
DB 1161 AACTTCTTCCAGAGACTTTTGAATATCATACATTCGGGTATATGATGAGAGGCAACG 1220
QY 387 GlnLeuLeuProHisTyrIlySglnIleThrHisArgPheIleGlnAlaAlaArgAlaGlnGly 406
DB 1221 GATCTCCCTGGCGTACTGGAATGACACTTACCAATTCATCTCTAAAGCAAGAACATGTA 1280
QY 407 ThrHisValleuValHisCysIlySmetGlyValSerArgSerAlaAlaThrValleuAla 426
DB 1281 TCTAATGCTTGTGTCACCTGCAAAATGGGGGAGTCCCTCAGCCTCCACCGCTGATGCC 1340
QY 427 TyrAlaMetIlySglnTyrGlnCysSerLeuGlnGlnAlaIleArgHisValGlnIleu 446
DB 1341 TATGCAATGAAGAAATATAGCTGATCTGACCGAGCTATGATGATGTGAAAGAAAGA 1400
QY 447 ArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeuGlnIleTyrGlnIlyIle 466
DB 1401 CGAAGGTATACCAACGCCACCAACCAAGCTTCATGACACAACTGGAAGATATCAGGGATC 1460
QY 467 LeuThrAla 469
DB 1461 TTGCTGGCA 1469

RESULT 8
US-09-775-925-1
; Sequence 1, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775.925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-775-925-1

Alignment Scores:
Pred. No.: 4.84e-98 Length: 1949
Score: 980.00 Matches: 183
Percent Similarity: 73.94% Conservative: 61
Best Local Similarity: 55.45% Mismatches: 84
Query Match: 40.23% Indels: 2
DB: 10 Gaps: 1
```

US-09-761-640-4 (1-471) x US-09-775-925-1 (1-1949)

```
OY 142 AspGluThrValLeuLeuGlyAlaSerPheProAspSerSerProSerGlyThrLeu 161
Db 32 GAGGAGAAATATCTCTCGAGTGAATCTTCCAGTAACGAAATGAAAGGCGACCAAT 91
OY 162 GlyLeuValLeuProLeuThrSerAspThrGlnValTyrLeuAspGlyAspGlyPhe 181
Db 92 GGGATGGTTCCTCCAGCTGGAGGACACGAAATCCACCTTGATGGAGATGTGGTTC 151
OY 182 SerValThrSerGlyGlnSerArgIlePheLysProIleSerIleGlnThrMetIrp 201
Db 152 AGCGTAGAGACAGACGAGATGCACATATTAAAGCCGTGTCTGTCCAGGCCATGTGG 211
OY 202 AlaThrLeuGlnValLeuHisGlnAlaCysGluAlaLeuGlySerGlyLeuValPro 221
Db 212 TCTGCCCTGCAGCTCTTCACAGGCTCTCGAAGTGGCCGAGGACCACTACTTCCCC 271
OY 222 GlyGlySerAlaLeuThrTrpAlaSerHisTyrGlnGluArgLeuAsnSerGlnSer 241
Db 272 GGGGGTAGAGCTCTCATCTGGGCTACCTACTGATGAGAGCTGCATCAGCTCCGAGCAG 331
OY 242 CysLeuAsnGluTrpThrAlaMetAlaAspLeuGluSerLeuArgPro-----ProSer 259
Db 332 TCCATCAACAGAGTGGAGACGCGCATGCGAGTCTACGCGGCCGCGCTCCCGCGG 391
OY 260 AlaGluProGlyGlySerSerGluGlnGluMetGluGlnAlaIleArgAlaGluLeu 279
Db 392 CTAATTGGGACACAGCCCACTGAGAGGAAAGGACCGAGCGCTCTATCAAGCCAGCTC 451
OY 280 TrpLysValLeuAspValSerAspLeuGluSerValThrSerLysGluIleArgGlnAla 299
Db 452 CCAACATCATGATGATGAGCCAGCATGACAAATGTCACTTCCAAAGATGCTAATGAA 511
OY 300 LeuGluLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPheIleAspAsnGlnMet 319
Db 512 TTAGAGAAACAGATGATTGTACTTGAAGAACTCAAGAAATTATTAACAATGAGAG 571
OY 320 LeuLeuLeuValAlaGlnIleArgAspArgAlaSerArgIlePheProHisLeuTyrLeuGly 339
Db 572 CTACTTATGTTGGAGACAGATGAGCAAGCCCTCTTATCTTGATCATCTTTATCTCGGC 631
OY 340 SerGluTrpAsnAlaAlaAsnLeuGluGluLeuGlnIleArgAsnArgValThrHisIleLeu 359
Db 632 TCTGAATGGAATGCATCCATCTGAGGAACTGACAGGCTCAGGGGTGATTACATTTA 691
OY 360 AsnMetAlaArgGluIleAspAsnPheTyrProGluArgPheThrTyrHisAsnValArg 379
Db 692 AATGTTACAGAGAAATCGATATATTTTCTGCTTATTTGCATATCATTAACATCGCA 751
OY 380 LeuTrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGluThrHisArgPheIle 399
Db 752 GCTTCAGATGAGAGACACACAGACCTCTCGCCCACTGGAATGAACCGATCATTTTAA 811
OY 400 GlnAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArg 419
Db 812 AACCAAGCGAAGAGAACCATTCCTCAAGTCCCTGGTCATTGCAAAATGGCGCGATCCGC 871
OY 420 SerAlaIleThrValLeuIleTyrAlaMetLysGlnTyrGluCysSerLeuGluGlnAla 439
Db 872 TCGGCTCCACAGCTTACCTATGCAATGAAAGAAATGGCTGGCTCTGGAAAAAGCA 931
OY 440 LeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGln 459
Db 932 TTTAACTATGTAAGAGAGAGGACACATACAGCGGCCCAAGCGGGCTTTATGAGGAG 991
OY 460 LeuGlnIleTyrGlnGlyIleLeuThrAla 469
Db 992 CTGTCTGAGTATGAGGACATCTTGATGCA 1021
```

RESULT 9
US-09-775-925-7
; Sequence 7, Application US/09775925

```
; Patent No. US20010049358A1  
; GENERAL INFORMATION:  
; APPLICANT: Lucbe, Ralf M.  
; APPLICANT: Wei, Bo  
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY  
; TITLE OF INVENTION: PHOSPHATASES  
; FILE REFERENCE: 200125.420  
; CURRENT APPLICATION NUMBER: US/09/775,925  
; CURRENT FILING DATE: 2001-02-01  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1052  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-775-925-7
```

```
Alignment Scores:  
Pred. No.: 1,94e-55 Length: 1052  
Score: 589.00 Matches: 107  
Percent Similarity: 75.35% Conservative: 55  
Best Local Similarity: 49.77% Mismatches: 45  
Query Match: 24.18% Indels: 8  
DB: 10 Gaps: 1
```

US-09-761-640-4 (1-471) x US-09-775-925-7 (1-1052)

```
OY 255 LeuArgProProSerAlaGluProGlyGlySerSerGluGlnGluMetGlnAla 274  
Db 190 CTCAGACCT-----ACTGAACGTGAAGACAGACAGAAAGGCTA 225  
OY 275 IleArgAlaGluLeuTrpLysValLeuAspValSerAspLeuGluSerValThrSerLys 294  
Db 226 ATTAAACCAAAATTGAAGGATCATGATGATGAGAGGATTTGGAAATATTCATCCAA 285  
OY 295 GlnIleArgGlnAlaLeuGlnLeuArgLeuGlyLeuProLeuGlnGlnIleTyrArgAspPhe 314  
Db 286 GAGATTAAGACAGAGCTTGGAAATGCAAAATGCTGCACACTTCGGGAAATCAAGAAATT 345  
OY 315 IleAspAsnGlnMetLeuLeuLeuValAlaGlnIleArgAspArgAlaSerArgIlePhePro 334  
Db 346 ATAGACAAATGAATGATGATGATCTTGGTCAAAATGATAGCCCTACACAGATATTGAG 405  
OY 335 HisLeuTyrLeuGlySerGluTrpAsnAlaAlaAsnLeuGluGluLeuGlnIleArgAsnArg 354  
Db 406 CATGCTTCTCGGGCTCAGAAATGGAATGCCCTCACTTGAAGAGCTTACAGAACGAGG 465  
OY 355 ValThrHisIleLeuAsnMetAlaArgGluIleAspAsnPheTyrProGluArgPheThr 374  
Db 466 GTACGGTATATCTTGAATGTCACGAGAGATGATTAATCTTCCACGAGCTTTGAG 525  
OY 375 TyrHisAsnValArgLeuTrpAspGluGluSerAlaGlnLeuLeuProHisTrpLysGln 394  
Db 526 TATCATTAACATTCGGGTATATGATGATGAGAGGACAGAGATCTCTGGCGTACGAAATGC 585  
OY 395 ThrHisArgPheIleGlnAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLys 414  
Db 586 ACTTCAAAATTCATCTCTAAACAAAGAAACATGATGCTTAATAGCTTGTGACATGCAAA 645  
OY 415 MetGlyValSerArgSerAlaIleThrValLeuAlaTyrAlaMetLysGlnIleTyrGlnCys 434  
Db 646 ATGGGGGTAGTGCCTCACACCTCCACCGTATTTGCTTATGCAATGAAAGAAATATGCTGG 705  
OY 435 SerLeuGlnGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnPro 454  
Db 706 AATCTGACCGAGCGCTATGATGATGATGAAAGAAAGAGCAAGCGTAAACCAAGCCCAACCA 765  
OY 455 GlyPheLeuArgGlnLeuGlnIleTyrGlnGlyIleLeuThrAla 469  
Db 766 AGCTTCATGAGACAACTGGAAGAGATATCAGGGAGTCTTCTGCGCA 810
```

RESULT 10
US-09-808-701-12

```
; Sequence 12, Application US/09808701
; Patent No. US20020146757A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Chen, Rui-hong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020146757A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2D
; CURRENT APPLICATION NUMBER: US/09/808,701
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PL_FL_genes Version 2.0
; SEQ ID NO 12
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(765)
US-09-808-701-12
Alignment Scores:
Pred. No.: 5.43e-50 Length: 969
Score: 539.00 Matches: 120
Percent Similarity: 59.55% Conservative: 39
Best Local Similarity: 44.94% Mismatches: 70
Query Match: 22.13% Indels: 38
DB: 10 Gaps: 7
US-09-761-640-4 (1-471) x US-09-808-701-12 (1-969)
QY 1 MetAlaLeuValThrValSerArgSerProGlySerGlyAlaSerThrProValGly 20
   |||||||
DB 34 ATGGCCCTGGTGCACCTGCAGCGCTGCCACGCCCGCGCTCTCTGGCCACG 93
   |||
QY 21 ProTPAsp-----GlnAlaValGlnArgSerArgLeuGlnArgGln 36
   :|:|:|
DB 94 AACAGCGAGTGTGAGGCTGGCAGCGAAGAGATCGAAATTAACTTCAGCTTAAGTGA 153
   :|:|:|
QY 37 SerPheAlaValLeuArgGlyAlaValLeuGlyLeuGlnAspGlyGlyAspAsnAsp 56
   ||||| :|:|:|:|
DB 154 ACCTTTTTCATGTGTAAGCGCAGCCCTCTTCTTACACAGAGGA----- 198
   |||
QY 57 AlaAlaGlnAlaSerSerIuProThrGluYsAlaProSerGluGluGluLeuHisGly 76
   ||| |||
DB 199 -----AGCAGCCCTCAA----- 210
   |||
QY 77 AspGlnThrAspPheGlyGlnGlySer---GlnSerProGlnIuYsGlnIu---GlnGln 94
   ||||| ||| ||| ||| |||
DB 211 -----GGCCAGCGGAGTCTTCAGCAACCCACACAGATCAGGTGATCG 255
   |||||
QY 95 ArgGlnHisLeuHisLeuMetValGlnLeuLeuArgProGlnAspAspIleArgLeuAla 114
   ||||| :|:|:|:|
DB 256 CCTCAACATCTTCAGGTGATGATCAACTTCTGCGTGCAGAGACAGATCAAGCTGCA 315
   |||||
QY 115 AlaGlnLeuGlnAlaProArgProProArgLeuArgTyrLeuValVala]-----Ser 132
   :|:|:|:|
DB 316 GTGGCGTGTGAAGCGCCCTGGGCGGAGCGGTCGGTGAATGATGATGATACACAC 375
   |||
QY 133 ThrArgGluGlyGluIuYsLeuSerGlnAspGluThrValLeuLeuGlyValAspPhePro 152
```

```
DB 376 GGGCCCGAG-----GACACCGAGGAGAAATCTTGTGAGTGCACCTTTCC 423
   |||:|:|
QY 153 AspSerSerSerProSerCysThrLeuGlyValLeuProLeuThrSerAspThrGln 172
   ||| |||||
DB 424 ACTAAGAAAGTAAAGAGCGACACATTGGATGGTCTCCAGTGGAGCCAGCAAA 483
   |||
QY 173 ValTyrLeuAspGlyAspGlyGlyPheSerValThrSerGlyGlyGlnSerArgIlePhe 192
   :|:|:|:|
DB 484 ATCCACCTTGATGAGATGCTGGTTACCGTCGACACAGCAAGAGATGCACATATT 543
   |||||
QY 193 LysProIleSerTleGlnThrMetThrAlaThrLeuGlnValLeuHisGlnAlaCysGlu 212
   |||||:|:|:|
DB 544 AAGCTGTGTCTGTCCAGGCCATGTGTTCTGCCCTGCAGGTCTTCACAAAGGCTGCGAA 603
   |||||
QY 213 AlaAlaLeuGlySerGlyLeuValProGlyIuYsSerAlaLeuThrThrAlaSerHisTyr 232
   |||||
DB 604 GTGGCCCGGAGGCACACACTTCCCGGGGCTGTAGCTTCATCTGGGCTACTACTAT 663
   |||
QY 233 GlnGluArgLeuAsnSerGluGlnSerCysLeuAsnGluThrThrAlaMetAlaAspLeu 252
   :|:|:|
DB 664 GAGAGCTGCATCAGCTCCGACGACAGCTGCATCAACGATGAGACGCCATGACGACCTG 723
   |||
QY 253 GluSerLeuArgProProSer 259
   |||||
DB 724 GAGTCTACGCGGCCGACTCC 744
   |||||
RESULT 11
US-09-815-419-3
; Sequence 3, Application US/09815419
; Patent No. US20020065406A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 18221, A NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE AND
; FILE REFERENCE: 10448-030001
; CURRENT APPLICATION NUMBER: US/09/815,419
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-815-419-3
Alignment Scores:
Pred. No.: 4.01e-19 Length: 654
Score: 256.00 Matches: 72
Percent Similarity: 41.00% Conservative: 35
Best Local Similarity: 27.59% Mismatches: 77
Query Match: 10.51% Indels: 77
DB: 10 Gaps: 5
US-09-761-640-4 (1-471) x US-09-815-419-3 (1-654)
QY 205 GlnValLeuHisGlnAlaCysGlnAlaAlaLeuGlySerGlyLeuValProGlyGlySer 224
   :|:|:|
DB 51 GAAGCAATGCACACAG-----GGTGAC 71
   |||
QY 225 AlaLeuThrThrAlaSerHisTyrGlnGluArgLeuAsnSerGluGlnSerCysLeuAsn 244
   ||| :|:|:|:|
DB 72 AACGCTAACTGGAAAGAAATTATAGAAACATGGAAGATGCCAATTCATGTTGT--- 128
   |||
QY 245 GluThrThrAlaMetAlaAspLeuGlnSerLeuArgProProSerAlaGluProGlyGly 264
   :|:|:|
DB 129 -----GGA 131
   |||
QY 265 SerSerGluGlnGluGlnMetGluGlnAlaIleArgAlaGluLeuThr-LysValLeuAs 284
   ||| ||||| ||| |||
```



```

; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/964,899
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-46

Alignment Scores:
Pred. No.: 1,99e-16 Length: 1337
Score: 235.50 Matches: 69
Percent Similarity: 45.80% Conservative: 40
Best Local Similarity: 28.99% Mismatches: 102
Query Match: 9.67% Indels: 27
Gaps: 6

US-09-761-640-4 (1-471) x US-09-964-899-46 (1-1337)

OY 232 TyrcIngluarleuanserserglInserCysleuasngluTrpThAlaMetAlasp 251
    ||| ||| |||
Db 594 TATGATGAGTATACCAATGAGCCAGCGAGTATGCCCTCCAGCCACTTACATAGTCTC 653
OY 252 leugluserleuarprproseralagluPro-----glyCyserserglu 267
    ||||| ||||| ||||| ||||| ||||| |||||
Db 654 CTCGAGTCTCGAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 713
OY 268 glnglIngluIngluIngluIngluIngluIngluIngluIngluIngluIngluInglu 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 714 TTTAAGCAGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 773
OY 285 ValSerAspLeuGluSerValThrSerLysgluLeuArgGlnAlaLeuGluLeuArgLeu 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 774 GTGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815
OY 305 glYleuProleuIngluIngluIngluIngluIngluIngluIngluIngluIngluInglu 324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 -----CCCATCCCAACCAACCCCTGAC--ATCGAGAACGCTGAGCTC----- 854
OY 325 glnarGAspArgAlaSerArgIlePheProHisLeuTyrluGluSerGluTrpAsnAla 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 -----ACCCCATCTTGCCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 899
OY 345 AlaAsnLeuGluGluLeuLeuArgAsnArgValThrHisLeuAsnMetAlaArgGlu 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 900 CAGGACCTGAGACACCATCATGAGCGCTGACATCGCTACGATCAACGTCACCATCAT 959
OY 365 IleAspAsnIleuTyrluGluArg-----PheThrTyrluHisAsnValArgLeuTrpAsp 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 960 CTTCCTCTACCATATAGAGAGCGCTGTCACTCAACGAGCGCTGAGCGCTGAGCGCTGAG 1019
OY 383 glugluseralagluLeuLeuProHisTrpLysgluTrpHisArgPheIleGluAlaAla 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 AGCAACAGCAGCAACCTCGGCGAGTCTTGAAGAGCGCTTGTGATTCATTCAGAGAACT 1079
OY 403 ArgAlaGlnGlyThrHisValLeuValHisCysLysMetGlyValSerArgSerAlaAla 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1080 CACCAAGTGTGGGAAGGCGCTTCTCATCTACATGCGCAGCGCTGCGGCTCCGCCACAC 1139
OY 423 ThrValLeuAlaTyrluAlaMetLysGlnTyrluGluCysSerLeuGluAlaLeuArgHis 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1140 ATCGCATCGCTTACTTGTATGAGACACACTGAGTACCATGACTGATTAATAATT 1199
OY 443 ValGlnGluLeuArgProIleAlaArgProAsnProGlyPheLeuArgGlnLeu 460
```

```

Db 1200 GTCAAGGCAAAAGCAACATATATCTCCCAAAACCTTACTTATGAGGAGGTG 1253
RESULT 14
US-09-878-574-1411
; Sequence 1411, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1411
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-038-Q1-B1-D7
US-09-878-574-1411

Alignment Scores:
Pred. No.: 1,39e-16 Length: 391
Score: 230.00 Matches: 52
Percent Similarity: 56.25% Conservative: 20
Best Local Similarity: 40.62% Mismatches: 56
Query Match: 9.44% Indels: 0
Gaps: 0

US-09-761-640-4 (1-471) x US-09-878-574-1411 (1-391)

OY 337 TyrluGlySerGluTrpAsnAlaAlaAsnLeuGluGluIngluIngluArgAsnArgValThr 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TATTTGGGTTCATTTGCCACCTGACCTTAACAGCCTGCTTAAAGACTGATATTAAT 60
OY 357 HisIleLeuAsnMetAlaArgGluIleAspAsnIleuTyrluGluArgPheThrHis 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CACATTTGACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120
OY 377 AsnValArgLeuTrpAsnGluGluSerAlaGlnLeuLeuProHisTrpLysgluThrHis 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ATTATGACGTGTGACAGAGATGATGAGACTTAAACAATCTTCAATGATGATGATGAT 180
OY 397 ArgPheIleGluAlaAlaArgAlaGlnGlyThrHisValLeuValHisCysLysMetGly 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GATTTATGATGAGAACCAAGAAAGACTCGGTGGAGGTTTGTGTTGATTTGTTGCTGGA 240
OY 417 ValSerArgSerAlaAlaThrValLeuAlaTyrluAlaMetLysGlnTyrluGluCysSerLeu 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGATCAAGAGAGTGTAGTATAGTGTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 300
OY 437 GluGlnAlaLeuArgHisValGlnGluLeuArgProIleAlaArgProAsnProGlyPhe 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 TTGAAGCTCTGAACAACAGTAAAGATAGACGACGACGACGACGACGACGACGACGACG 360
OY 457 LeuArgGlnLeuGlnIleTyrluGln 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ATTTGTCACTGAGAGACTTTGAA 384

RESULT 15
US-09-816-494-3
; Sequence 3, Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 And 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
```

```

? CURRENT APPLICATION NUMBER: US-09/816-494
? CURRENT FILING DATE: 2001-03-23
? PRIOR APPLICATION NUMBER: US 60/191,858
? PRIOR FILING DATE: 2000-03-24
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 1998
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-816-494-3

```

Alignment Scores:	
Pred. No.:	2,67e-15
Score:	227.50
Percent Similarity:	58.39%
Best Local Similarity:	39,42%
Query Match:	9,34%
DB:	10
Gaps:	2
Length:	1998
Matches:	54
Conservative:	26
Mismatches:	54
Indels:	3
Gaps:	2

US-09-761-640-4 (1-471) x US-09-816-494-3 (1-1998)

[illegible]

```
Search completed: April 12, 2003, 13:57:40
Job time : 371 secs
```